

SOCIOMETRY

A JOURNAL OF RESEARCH IN SOCIAL PSYCHOLOGY

VOLUME XIX • NUMBER 3 • SEPTEMBER 1956

CONTENTS

Group Norms among Bomber Crews: Patterns of Perceived Crew Attitudes, "Actual" Crew Attitudes, and Crew Liking Related to Aircrew Effectiveness in Far Eastern Combat. . . *Leonard Berkowitz* 141

Status Symbols and the Measurement of Status Perception
James A. Davis 154

Patterns of Perceived Interpersonal Relations
Rosalind Dymond Cartwright, Julius Seeman and Donald L. Grummon 166

Cognitive Similarity in Facilitating Communication
Philip J. Runkel 178

The Role of Test Intelligence and Occupational Background as Factors in Occupational Choice. *C. T. Pihlblad and C. L. Gregory* 192

SOCIOMETRY

EDITORIAL POLICY

Sociometry is concerned with the entire range of interests and problems represented by research in social psychology. It is the policy of the editors to seek those manuscripts for publication which represent the significant research interests of investigators who are concerned with giving the field of social psychology theoretical structure and reporting research which is clearly focused, well designed, and competently conducted.

While social psychology is presently regarded by most as a field with indeterminate boundaries, it has as its central focus the investigation of the processes and products of social interaction at the interpersonal, intrapersonal, intergroup and intragroup levels and the development of significant generalizations therefrom. In keeping with the more general meaning of the name of the journal emphasis will be placed on measurement of social behavior. However, this emphasis does not exclude the acceptability of good articles which must rely upon qualitative materials and analyses.

The editors and editorial consultants can be expected to subject manuscripts to rigorous criticism and screening according to the best standards of scientific research and at the same time avoid a sterile orthodoxy which would stultify the communication of creative ideas at the growing edge of the science. Thus the journal will strive to be flexible in its response to the publication needs of its contributors.

It is the intention of the editors to avoid any tendency toward professional provincialism and to invite contributions from any sector of the scientific community which promise to further the objectives of the journal.

Annual subscription rate \$9.00; special rate to members of THE AMERICAN SOCIOLOGICAL SOCIETY \$4.50. Single issues \$2.50.

Specifications for Manuscript Preparation for Sociometry may be obtained from the Editorial Office.

Back Issues dated 1937 through 1955 should be ordered from the former publisher, Beacon House, Beacon, New York.

Notice of change of subscribers's address should be sent to THE SOCIETY four weeks in advance, indicating the old as well as the new address.

Group Norms among Bomber Crews: Patterns of Perceived Crew Attitudes, "Actual" Crew Attitudes, and Crew Liking Related to Aircrew Effectiveness in Far Eastern Combat¹

LEONARD BERKOWITZ, *University of Wisconsin*

Various studies and theoretical works in recent years have called attention to the importance of group norms in influencing the behavior of group members (e.g., 9, 10). Basic to most definitions of a norm is the idea of behavioral similarity stemming from participation in the group and reinforced by the perception that behavioral conformity is necessary in order to maintain one's membership in the group. A series of laboratory experiments (4) has demonstrated that individuals will tend to conform to the perceived attitudes of others in their group if group membership is attractive to them, i.e., if the group is cohesive. Defining group norms in terms of the pattern of perceived group attitudes and group cohesiveness, Berkowitz (2) showed that individuals will conform to a group norm for high production (perceived high task motivation among the group members and high cohesiveness) or to a group norm for low production (perceived low task motivation among the members and high cohesiveness) even when there is no overt pressure placed upon them for conformity.

The general hypothesis behind several studies carried on by the Crew Research Laboratory is that aircrews which have developed group norms favorable to Air Force objectives will tend to be more effective crews. Within the framework of these studies, we may say that a group norm to be an effective crew exists when the crew members perceive each other as being motivated to achieve crew effectiveness and when crew membership is attractive to them. One factor which may make for cohesiveness in most groups, and which is relatively easily measured, is the liking of group members for each other. A group norm for crew effectiveness thus is defined operationally in terms of the combination of high crew liking and perceived high task motivation. Crews exhibiting this pattern should receive relatively high superiors' ratings of crew performance and should have relatively few incidents of task-avoidance behavior. The superiors' ratings should reflect not only better group performance—to the extent that group striving for effectiveness leads to effectiveness—but also favorable superiors' reactions to the crew's "eagerness," i.e., behavior in conformity

¹ This study was carried out at the Crew Research Laboratory of the Air Force Personnel and Training Research Center. The opinions expressed in this paper are those of the author and are not necessarily the views of the United States Air Force.

to the crew norm. Group norms for crew effectiveness should also be reflected in a smaller number of task-avoidance behaviors such as aborted missions.

In the first of two preceding studies in this series (1), the category of crews defined at the completion of the three-months training course as being highly cohesive and in which the members perceived each other as being highly task motivated contained significantly fewer crews with aborted missions than the categories composed of the other combinations of cohesiveness and perceived crew motivation. Similar results were obtained in the second study employing perceived motivation data collected shortly after the complete crew had been assembled for training (3). There were fewer aborts throughout training when the crew members in the early part of their training had relatively high liking for each other and perceived that the crew was highly motivated than when this perception existed but the crew members had less liking for each other. The crews in the former condition also tended to be rated higher by their instructors than the crews in the latter condition. These results, furthermore, could not be attributed to the "actual" mean level of motivation for the individual crew members as determined by their responses to an attitude questionnaire.

The present study was designed to replicate the earlier findings employing a different population of aircrews. An attempt was made to determine whether the effectiveness of bomber crews which had been in existence a relatively long time and were engaged in combat varied with different patterns of perceived crew motivation, actual crew motivation, and crew liking.

PROCEDURE

Subjects

There are 81 eleven-man B-29 crews in three medium bombardment wings of the Far Eastern Air Force in the present sample. The majority of these crews had been together, and relatively intact, for about nine months from the time of their assembly at Randolph Air Force Base, Texas, and their median time in this theater of operations was four months. The crews, regularly flying combat missions over Korea, were stationed in Japan and Okinawa. No crew was included in this sample if attitudinal data were lacking for more than three of the eleven crew members. The majority of crews were complete, and a large number of the remaining crews had only one member missing.

The Attitude and Perceived Attitude Measures

The present data were collected by a team of psychologists who had gone overseas to Japan and Okinawa in the spring of 1953. The attitudinal

data employed in this study were obtained from the aircrew members' responses to an attitude questionnaire developed at the Crew Research Laboratory (7).

Perceived motivation. One of the items from this questionnaire was utilized in the perceived motivation measure. The item was: "The B-29 is an important part of the defense of the United States."²

The item was presented on a separate sheet, together with the following instructions:

"We would like you to indicate how many members of your crew you believe would agree with the following statement and how many would disagree with it. The total is 10 (11 minus yourself)." The crew scores used were simply the total sum of crew members perceived as having *favorable* attitudes, i.e., perceived as agreeing with the statement.

If the perceived motivation item was to assess the presence of a group norm, responses to this item should be affected by within-group conditions. In other words, within-crew variance on item responses should be significantly less than between-crew variance. An analysis of variance was undertaken employing the data from 48 complete crews. The Hoyt r (6) was .34 ($P = .02$) thus satisfying this criterion.

Since there were varying numbers of members responding to the item in the sampled crews, the group measure employed was the percentage of the maximum possible score. Thus, if each of the eleven crew members predicted that everyone else would respond favorably to the item, the crew would have a score of 110 out of the total possible of 110, or 100 per cent.

"Actual" motivation. Responses to this same item in the main body of the attitude questionnaire were used in arriving at the crews' "actual" motivation scores. It is assumed that crew task motivation is higher the greater the number of crew members actually agreeing with the statement that the B-29 is important to the defense of the United States. For every crew we obtained the total number of crew members actually agreeing³

² The content of this item suggests that it has to do with the crew members' acceptance of their particular role in the Air Force. In using this item as a measure of task motivation we are assuming that crew members who say the B-29 is important to the defense of the United States also feel that it is important to have an effective B-29 crew, and that this attitude is highly correlated with *wanting* to have an effective crew. There is evidence supporting these assumptions. Other items in the "motivation" scale from which this item was taken include: "I am willing to do all I can to make my crew a really good bomber crew," and "I am very anxious to have a really good bomber crew."

³ "Agreement" is here defined in item analysis terms. Each of the items was scored dichotomously in that the five original alternatives to each item were reduced at the time of scoring to two, with these two responses dividing the population as nearly in half as possible. The alternative chosen by the respondents with a high score on the total scale was defined as "agreement."

with this Likert-type item and divided this sum by the number of crew member respondents. These "actual" motivation scores correlated .49 with the perceived motivation scores for all the crews in the sample.

Crew cohesiveness was measured by a six-item Likert-type scale designed to get at the liking of the crew members for each other. The two items in the scale with the highest correlations with the total scale score are: "The members of my crew are the kind of people I like to spend a lot of time with," and "Everyone has a lot of respect for everyone else on my crew."

The Criteria⁴

There are two criteria of aircrew effectiveness in the present study: superiors' ratings of over-all crew effectiveness and a measure of the frequency of aborted, or failed, combat missions.

The superiors' ratings are taken in this study to be a relatively global measure of aircrew effectiveness. The abort measure, at one level of abstraction, has to do with avoidance of the aircrew members' assigned task: flying to and bombing the primary target.

1. *Superiors' ratings of crew effectiveness.* This measure is based on ratings by the crews' superiors (generally two squadron officers—the commanding officer and operations officer—and two wing officers) on over-all crew effectiveness. The Horst reliability (5) for this measure based upon 77 crews was .78.

2. *The percentage of assigned missions which failed (or were aborted).* A score on this criterion could be due to any one of the following:

- a. A crew abort with or without a replaced aircraft.
- b. Trouble after take-off that prevented the plane from reaching the target.
- c. Bombing a secondary target rather than the assigned primary target.
- d. Failure of the crew to drop all its bombs on the primary target.

The reliability of this measure was ascertained by splitting the number of assigned missions in half (i.e., early and late in tour, relatively speaking) and interrelating the percentage of failure events in each half. This Pearson r , based on 88 crews, is .47. The number of aborted missions was adjusted for differences in the length of the tour by dividing the frequency of aborted missions by the number of missions assigned to the crews. The percentage scores were normalized for the present study by means of an arc sine transformation.

The Statistical Analyses

We are mainly concerned in the present study with interactions among the three attitudinal variables: perceived crew motivation, actual crew

⁴ Dr. D. G. Forgays, of the Crew Research Laboratory, analyzed the criterion data and provided the criterion measures for the present study.

motivation, and crew liking. Analysis of variance techniques were employed to test the significance of these interactions. The distributions on each of the three variables were dichotomized as close to the median as possible, and the sample of crews was divided into the eight possible combinations of "high" and "low" scores on the variables.

It is apparent from the figures shown in Table 1 that the variables are interrelated—the conditions with the greatest number of crews are the high perceived motivation, high actual motivation, high cohesiveness (high, high, high) condition and the low perceived motivation, low actual motivation, low cohesiveness (low, low, low) condition. The resulting disproportionality in the cell frequencies ruled out the possibility of employing the conventional analysis of variance procedures. Instead, the methods suggested by Snedecor (13) were followed.

Indirect methods were employed in testing the interactions. This was done in three stages.⁵ First, each variable was ignored in turn to test the first-order interactions between the other two variables. This was a 2×2 design assuming the presence of interactions (13, pp. 287-288). In the second stage we determined the interactions between each of the three variables taken singly and the four combinations of the other two variables. This was a 4×2 design, also assuming the presence of interactions (13, pp. 291-293). This last analysis is analogous to determining how the scores in four different conditions are affected by a contingency variable. Third, *t*-tests were run between the scores for selected pairs of conditions. This last procedure, based upon suggestions by Lindquist (8, pp. 90-96), employs the within-conditions mean square based upon all eight conditions in the error term in the *t*-test. For each of the criteria preliminary analyses of variance were first computed. The within-conditions mean squares were obtained from these preliminary analyses.

RESULTS

Superiors' Ratings of Over-all Crew Effectiveness

Table 1 presents the means for superiors' ratings in the eight conditions. Most of the previous investigations in this area (e.g., 2, 11) have paid little if any attention to individual characteristics that may contribute to the

⁵ Briefly, these procedures employ the differences between condition means with each of these means weighted not by the number of cases in the condition—this might provide a bias in favor of some conditions—but by the harmonic mean of all the cell frequencies. There are eight condition means, and these means are compared in various ways depending upon the variables being studied. Three estimates of each main effect are obtained. Because it is the condition means rather than the condition sums that are compared there is some variation in the resulting main effect mean squares. In general, however, these results are in agreement with each other. Where one estimate attains statistical significance the other two estimates of this main effect are also significant.

TABLE 1

Mean Superiors' Ratings of Over-all Crew Effectiveness in the Eight Conditions

Conditions	Actual Motivation	
	High (A)	Low (B)
1. High liking, high perceived motivation	(15)* 72.6	(8) 70.1
2. High liking, low perceived motivation	(5) 63.2	(8) 65.2
3. Low liking, high perceived motivation	(8) 70.2	(11) 66.4
4. Low liking, low perceived motivation	(8) 64.5	(16) 71.4

* The figures in parentheses are the cell frequencies.

TABLE 2

Summary of Mean Squares: Superiors' Ratings*

Test	Main Effects			Interactions	
	Perceived Motivation	Actual Motivation	Crew Liking	Two \times two (DF = 1)	Four \times two (DF = 3)
A. <i>Two \times two</i>					
1. Crew liking \times perceived motivation	173.17		3.72	320.64†	
2. Perceived motivation \times actual motivation	196.63	10.43		385.43‡	
3. Crew liking \times actual motivation		1.13	5.66	95.48	
B. <i>Four \times two</i>					
4. "Conditions" \times perceived motivation	232.16				203.01
5. "Conditions" \times crew liking			2.88		100.19
6. "Conditions" \times actual motivation		18.28			119.10
Within Conditions: 86.40					

* Cf. footnote 5 for an explanation of the differences in main effects mean squares.

† *F* ratio of mean square in question over "within" mean square significant at .10 level of confidence.‡ *F* ratio of mean square in question over "within" mean square significant at .05 level of confidence.

group's influence over the individual members. In this study we have made an initial attempt to assess whether the actual motivation of the crew members interacts with the combinations of crew liking and the perceived crew motivation. The data for these four combinations are therefore presented separately for the crews with high actual motivation and for those with lower motivation.

A summary of all the analyses of variance for this criterion is given in Table 2. None of the main effects are statistically significant, but there are two first-order interactions that are either significant or nearly so: perceived

motivation and actual motivation, and perceived motivation and crew liking. (These interactions are computed from the 2×2 tables. That is, the analyses are based on comparisons of means in four conditions—high and low on each of two variables with the third variable collapsed.)

The interaction of perceived motivation and crew liking is as predicted by the hypothesis. The highest rated crews perceive high motivation in themselves and are relatively cohesive while the lowest rated crews are also relatively cohesive but perceive lower task motivation. Under the condition of high cohesiveness, then, the perceived group motivation is related to the ratings of group performance.

The interaction of perceived and actual motivation was not predicted. In this analysis the lowest rated crews are those whose actual motivation is high, but who perceive low task motivation in themselves. The best crews have high actual and perceived motivation. Over all the crews in the sample, the poorer crews tend to have a discrepancy between perceived and actual crew motivation.

The effects associated with the third or collapsed variable are not shown in the above two-by-two analyses. We may gain some insight into the interactions among all three of the attitudinal variables in the comparisons of the means for the eight separate conditions as given in Table 1.

The above-mentioned two first-order interactions—(a) the predicted interaction of cohesiveness and perceived crew motivation, and (b) the unexpected interaction of perceived and actual crew motivation—will be taken as the starting point in organizing the obtained results. An attempt will be made to determine whether these interactions refer to relatively distinct psychological phenomena, or whether the significance of one of the interactions can be traced to the effects of the other interaction.

The Interaction of Perceived and Actual Crew Motivation

Inspection of Table 1 reveals that crews in the low, low, low condition (4B) tend to have a relatively high superiors' rating. This is surprising unless the possible negative effects of a discrepancy between perceived and actual motivation are considered. In condition 4B there is a relatively small discrepancy between these measures—both scores are below their respective means. The crews in condition 3A also have a relatively small discrepancy between these measures—in this case both means are high—and these crews also tend to receive high superiors' ratings. The mean rating given to the crews in condition 3B is of the same order as the mean rating assigned to the crews in 4A, and in both of these conditions there is a larger discrepancy between the perceived and actual motivation measures. A two-by-two analysis of variance computed for these four low cohesiveness conditions yields an interaction significant between the 10 and 5 per cent

levels of confidence. (The within-conditions mean square based on all eight conditions was used as the error term.)

The Interaction of Cohesiveness and Perceived Crew Motivation

The above four conditions are all characterized by low crew cohesiveness. An examination of Table 1 suggests that the magnitude of the discrepancy between the two motivation measures is not related to superiors' ratings for the high cohesiveness conditions. Condition 2B, for example, has a small discrepancy between perceived and actual motivation—both scores are relatively low—but unlike the crews in the comparable condition, 4B, these crews are rated low by their superiors. The obtained results do suggest, however, that it is the level of perceived crew motivation that is related to superiors' ratings in the high cohesiveness condition. In general, the higher the perceived motivation, the better the crew. However, actual motivation must also be high if perceived motivation is to be significantly associated with high ratings. There is a near significant difference ($P = .06$) between the mean ratings of the crews with high perceived motivation as compared with crews with low perceived motivation only for those crews whose actual motivation is also high. It may be that the members of cohesive crews were influenced to conform to the perceived crew attitude and perform the behaviors measured by the superiors' ratings only when their own predispositions were also in this direction. For this criterion, then, the actual level of crew members' motivation may make a difference in determining the extent of the effect of the group norm. It is obvious, however, that better controlled investigations are necessary to test this possibility.

Percentage of All Assigned Missions Which Failed

The mean percentage of failed missions (transformed into arc sines) for each of the eight conditions is given in Table 3. Table 4 presents the summary of all the analyses of variance with this criterion. As is shown in Table 4, there is a significant main effect: that for perceived motivation. Those crews with high perceived motivation scores have significantly fewer failed missions than the crews with relatively low scores on this measure. All the two-by-two and four-by-two interactions are also either significant or nearly so. The present variables, therefore, all interact in relating to the criterion of aborted combat missions. However, in order to determine the manner in which these variables interact, comparisons must be made between condition means as was done with the superiors' ratings criterion.

The differences between condition means, as shown in Table 3, are similar to the differences between the condition means for superiors' ratings except that here a high score indicates "poor" performance. As with su-

TABLE 3

Mean Per Cent of Failed Missions in the Eight Conditions (Arc Sine Transformation)

Conditions	Actual Motivation	
	High (A)	Low (B)
1. High liking, high perceived motivation	(15) 9.49	(9) 15.42
2. High liking, low perceived motivation	(5) 22.94	(8) 26.91
3. Low liking, high perceived motivation	(7) 11.69	(9) 19.39
4. Low liking, low perceived motivation	(8) 25.52	(14) 13.14

TABLE 4

Summary of Mean Squares: Per Cent of Failed Missions*

Test	Main Effects			Interactions	
	Perceived Motivation	Actual Motivation	Crew Liking	Two \times two (DF = 1)	Four \times two (DF = 3)
A. Two \times two					
1. Crew liking \times perceived motivation	1031.40†		52.30	640.36†	
2. Perceived motivation \times actual motivation	1018.75†	3.23		828.62†	
3. Crew liking \times actual motivation		91.98	4.38	598.24†	
B. Four \times two					
4. "Conditions" \times perceived motivation	1110.88†				458.89†
5. "Conditions" \times crew liking			26.24		338.05§
6. "Conditions" \times actual motivation		28.00			416.84†
Within Conditions: 136.35					

* Cf. footnote 5 for an explanation of the differences in main effects mean squares.

† *F* ratio significant at .01 level of confidence.‡ *F* ratio significant at .05 level of confidence.§ *F* ratio significant at .10 level of confidence.

periors' ratings, the crews in the low, low, low condition (4B) tend to perform fairly effectively—they have relatively few failed missions. This finding will again serve as a starting point in the discussion, and we can apply the same interpretations to the present data as were applied to the relationships with superiors' ratings.

In the low liking conditions (conditions 3 and 4) the discrepancy between perceived and actual motivation is related to the percentage of failed missions regardless of whether these motivation scores are both high or low. A two-by-two analysis of variance of the means in these conditions, using

the total within conditions mean squares as the error term, reveals a statistically significant interaction between the perceived and actual motivation measures ($F = 6.56$, $p = .05$). The greater the discrepancy between perceived and actual motivation when there is little liking among the crew members, the greater the percentage of failed missions.

In the high cohesiveness conditions (1 and 2) there are significantly fewer failed missions ($p = .05$) for crews with high perceived motivation than for crews with low perceived motivation. Unlike the case for superiors' ratings, however, this difference is significant for both the crews with high and low actual motivation. Within the range of task motivations represented by the present data, the members of the cohesive crews apparently tend to conform to the perceived crew attitude.

DISCUSSION

It is, of course, impossible to determine the direction of the casual relationships in correlational studies such as this. There is a good possibility, for example, that the crews' scores on the three attitudinal variables—perceived and actual motivation and cohesiveness—were affected by past performance effectiveness. However, we then must ask why high scores on these variables are associated with a high level of effectiveness for some crews and with a comparatively low level of effectiveness for other crews. Some crews with high cohesiveness have relatively few failed missions while other crews high in cohesiveness have relatively many failed missions. A similar phenomenon can be described for the other measures. If these attitudinal variables *are* affected by crew performance, there must be psychological characteristics of the crews (contingency conditions) which determine the degree and direction of the relationship between the attitudinal variables and performance. Furthermore, these contingency conditions may well be described in terms of the level of the present attitudinal variables. In other words, patterns of the present attitudinal variables may both affect and then be affected by crew performance.

The most obvious illustration of this possibly reciprocal relationship can be found with crew norms. Crews performing well, as defined by the Air Force, may be motivated to do so because the members perceive the others in their attractive crews as desiring this mode of behavior. The combination of perceived crew motivation and crew cohesiveness causes this behavior. But then, after several missions in which the crew performs effectively, crew members truthfully can describe the crew as being highly motivated. Moreover, since the crew's performance is in conformity with the norm, the crew cohesiveness should be heightened.

This reciprocal relationship is only conjecture. There is little doubt, however, that group norms of the type measured here can antedate and affect the groups' performance (2, 3, 12).

The Interaction of Perceived and Actual Motivation

Several hypotheses are suggested by the significant interactions between perceived and actual crew motivation in the low liking conditions. For example, conflict or intracrew tension may be related to discrepancies between these measures, and this tension could have resulted in lower superiors' ratings and the increased proportion of failed missions. Furthermore, the intracrew tension may be most likely to arise as a result of the misperceptions when, as is true in the present case, the crews are relatively less cohesive.

The Interaction of Cohesiveness and Perceived Motivation

If the previous analysis is correct, the effects or concomitants of discrepancies between perceived and actual motivation may affect performance in the low cohesiveness conditions but not in the high cohesiveness ones. In the high cohesiveness conditions, then, the level of perceived motivation within the crew should be directly related to the effectiveness of the crew performance.

The results in general support this hypothesis. When crew cohesiveness is high, there tend to be fewer failed missions and higher superiors' ratings of crew effectiveness for crews with high perceived motivation than for crews with low perceived motivation. The members of the cohesive crews may have conformed to the perceived crew motivation. This is true apparently for the abort criterion regardless of the actual level of motivation of the crew members, although there is a slight difference in the means for the high and low motivated conditions (A and B). For superiors' ratings the crews with high perceived motivation have significantly higher ratings than the crews with low perceived motivation when their actual motivation was also relatively high. This difference is not significant for the crews with lower actual motivation.

In neither the present nor the earlier studies in which somewhat similar results were obtained (1, 3) do we have any direct evidence that there is influence not to abort missions when crew liking is high and the crew is perceived as being motivated to be an effective crew, or that crew members are overtly influenced to abort the mission when the crew norm is opposed to this designated task. We may hypothesize that there is a "twilight zone" of difficulty, in which it is possible either to decide to fly to the primary target or not to do so, and that the group norm may affect the decision at this choice point.

The crews with a norm for effective performance, as defined in this study, tend to perform better—but not significantly so—than the crews with low cohesiveness and small discrepancies between perceived and actual motivation. It may be that the members of these latter crews have learned to adjust to each other so that they each carry out their assigned tasks with

a minimum of disruption. However, when group norms exist, as here defined, those crews with norms for effective performance apparently perform better than those crews with norms somewhat opposed to effective performance.

SUMMARY

Measures of aircrew members' perceptions of the task-oriented motivation of their fellow crew members were obtained for crews in combat in the Far Eastern Air Force. In addition, attitude items were administered measuring the crew members' "actual" motivation to perform their designated tasks and the liking of the crew members for each other. Criteria of crew effectiveness were also obtained, based on either superiors' ratings or the incidence of task-avoidance behaviors (the percentage of assigned missions in which all the crew's bombs were not dropped at the primary target).

The relationships between the attitude measures on the one hand, and criteria of superiors' ratings and percentage of failed missions on the other, tended to be similar and statistically significant. Based on these criteria the following conclusions were drawn:

Two different sets of psychological processes may have been operating to bring about the present results.

1. Under conditions of high liking among the crew members (high cohesiveness) the crew members may have been motivated to conform to the perceived group norm. Therefore, the better crews in the high cohesiveness condition were those who had a group norm to perform their designated tasks effectively, i.e., the perceived motivation among the crew members was high.

2. Under conditions of low cohesiveness, however, it is the magnitude of the discrepancy between actual and perceived crew motivation that was related to the criteria. The better crews were those for whom there was a relatively small discrepancy between these measures regardless of whether the scores were both high or both low.

Manuscript received: February 24, 1956

Leonard Berkowitz
Department of Psychology
University of Wisconsin
Madison 6, Wisconsin

BIBLIOGRAPHY

1. Berkowitz, L., *Group Norms and the Readiness of Aircrews to Complete Training Missions*, San Antonio, Texas: Human Resources Research Center, Lackland Air Force Base, December, 1952 (Research Note CCT 52-2.).
2. Berkowitz, L., "Group Standards, Cohesiveness, and Productivity," *Human Relations*, 1954, 7, 509-519.

3. Berkowitz, L., *The Perception of Group Attitudes as Related to Criteria of Group Effectiveness*, San Antonio, Texas: Air Force Personnel and Training Research Center, Lackland Air Force Base, November, 1954 (*Research Bulletin*, 54-62.).
4. Festinger, L., "Informal Social Communication," *Psychological Review*, 1950, 57, 271-282.
5. Horst, P., "A Generalized Expression for the Reliability of Measures," *Psychometrika*, 1949, 14, 21-31.
6. Hoyt, C., "Test Reliability Obtained by Analysis of Variance," *Psychometrika*, 1941, 6, 153-150.
7. Knoell, Dorothy M., and G. Stice, *Development of Attitude Scales Relevant to Combat Crew Membership*, San Antonio, Texas: Air Force Personnel and Training Research Center, Lackland Air Force Base, November, 1954 (*Research Bulletin*, 54-63.).
8. Lindquist, E. G., *Design and Analysis of Experiments in Psychology and Education*, Boston: Houghton Mifflin, 1953.
9. Newcomb, T. M., *Social Psychology*, New York: Dryden Press, 1950.
10. Roethlisberger, F. J., *Management and Morale*, Cambridge, Mass.: Harvard University Press, 1950.
11. Schachter, S., "Deviation, Rejection, and Communication," *Journal of Abnormal and Social Psychology*, 1951, 46, 190-207.
12. Schachter, S., N. Ellertson, D. McBride, and D. Gregory, "An Experimental Study of Cohesiveness and Productivity," *Human Relations*, 1951, 4, 229-238.
13. Snedecor, G. W., *Statistical Methods* (4th Ed.), Ames, Iowa: Collegiate Press, 1946.

Status Symbols and the Measurement of Status Perception¹

JAMES A. DAVIS, *Harvard University*

This analysis describes an attempt to use status symbols (a) to test the unidimensionality of perceived social prestige; (b) to develop an instrument for measuring subjective status; and (c) to measure ability to judge prestige. The symbols in the study are photographs of living rooms, the subjects are housewives in Cambridge, Massachusetts, and the measurement model is the Guttman scale.

THE RESEARCH PROBLEM

Subjective status may be defined as a person's belief about his location in a status order. This belief may or may not be congruent with his objective status, the status accorded to him by others. In any institutionalized status system, by definition, subjective and objective status should be correlated, but the degree of correlation is an empirical problem.

Measurement in the field of stratification has concentrated on the development of measures of objective status or external measures which purport to predict subjective status (10). Those studies which have attempted to measure subjective status directly have tended to beg the question of unidimensionality. Centers (2) provided his subjects with an *a priori* ranked list of classes with which to identify. Hetzler (8) proposed to avoid this problem by giving his subjects a graphic continuum, but his technique assumed the existence of an underlying dimension. Methodologically, it would seem that the fairest test would be to present a series of unordered "items" and use a measurement model such as the Guttman scale or latent structure analysis to test for unidimensionality.

What then shall be the items? Subsequent research using the Centers items (5, 10) has cast doubt on the generality of perception of nameable class strata in urban communities. The use of individual persons as "items" limits research to small communities and to only one such community. Material status symbols present themselves as an alternative for methodological and theoretical reasons. First, they may be used as discrete

¹ This report is drawn from the writer's doctoral dissertation, "Living Rooms as Symbols of Status: A Study in Social Judgment" (Harvard University, 1955). Parts of the study were reported in a paper read at the annual meetings of the American Sociological Society, September, 1955. I would like to acknowledge the aid of Professor Peter Rossi, my thesis advisor; my wife, Martha Davis, who conducted the majority of the interviews in the study; and Dr. Joseph A. Kahl.

items (e.g., specific living rooms, houses, automobiles, or items of clothing). Second, they are "manifest" in Lazarsfeld's sense of the term and do not force an order for semantic reasons. Third, the American merchandising system ensures that the same possessions will be known to people in many communities and status positions. Fourth, considerable research, particularly that of Chapin (3), shows material possessions to be an objective correlate of status. Hence, any lack of subjective perception would not be attributed to asking questions irrelevant to status. Fifth, a considerable body of literature, from Veblen (17) to Lynes (11) has stressed the importance of material symbols in the American status system, a phenomenon often slighted in sociological research.

To test this approach, the writer collected a series of photographs of contemporary American living rooms. They were largely gathered from national magazines and the files of public housing agencies. Sampling was impossible, but an attempt was made to cover a broad range of social status. After a pretest, 24 photographs were selected for the study. Interviews were conducted during the winter of 1954 with a probability sample of 134 Cambridge, Massachusetts, housewives, stratified by age and prestige of husband's occupation. The subjects were told that the study concerned "first impressions." After certain preliminary questions, each was asked to sort the pictures into four piles. Sizes of the piles were not specified. The sorting was to be done in terms of "the social standing of the people who live in the homes." No further definitions of status or social standing were given. When this was completed, each was then asked to choose a pile "most like your own in social standing" and then to sort the pictures in that pile into two groups: those which were "a little higher" than her own position and those which were "a little lower." Thus, each subject presented both a status distribution of the pictures and a dichotomization into pictures higher and lower than her own position.

Unidimensionality of Perceived Prestige

Five of the women refused to sort the pictures. Three were low-status, older women, all foreign born, who claimed they were unable to do the task, the comment of subject 61 being typical:

There are rich and poor, but I cannot tell. I do not know any people, so I have not seen houses. I would not know what to say. I have spent whole life on children. . . .

One was a European-born, middle-status subject who said that she had not been in this country long enough to learn the status system and that her sorting would reflect European standards. One was an exceedingly high-status woman who said:

Now that everyone has money, you can't tell a thing about people from their furnishings.

Two characteristics of these "refusals" are of interest. None of the five disclaimed the existence of a stratification system in this country, and all were in the older age groups.

Guttman-scale analysis of the remaining 129 protocols was done as follows. The sample was divided into random halves and scales were constructed on one half and replicated on the other. Item-to-total contingency tables indicated that only one picture would not scale, and this picture (the only one which had shown significant class differences in ratings in the pretest) was excluded from the scaling. Many subjects had insisted that a few of the pictures were neither "higher" nor "lower" but the "same." Therefore, our response system was turned into a trichotomy, and the problem of cutting points arose. Because we were interested in the order of items as well as subjects, we made two different scales, each with the same cutting point for all 23 pictures.

After this analysis it was possible to replicate the scales as a paper-and-pencil test on a group of 37 female students at the 1954 summer session of Rockford College, Rockford, Illinois. All the subjects were schoolteachers between the ages of twenty-one and sixty-five.

The results of these scaling operations are presented in Table 1.

With the exceptions of G1 and G2 in the Rockford sample, which are .01 and .03 below the conventional .90, all the scales meet the criterion of reproducibility. Furthermore, the scale types gave a satisfactory marginal distribution. For one picture the marginal was 0 per cent. No subject said that the lowest status picture was "higher" or "same." This gave us artificially inflated reproducibility, but recomputing reproducibility on half

TABLE 1
Reproducibilities for Various Scales

Scale	Number of Items	Sample		
		1	2	Rockford
G1	23	.909	.900	.890
G2	23	.910	.899	.873
H1	4	.985	.992	.972
H2	9	.955	.945	.905

Note: G1 is a Guttman scale. For all items a favorable response was "higher" and an unfavorable response was "same" or "lower." G2 is a Guttman scale. For all items a favorable response was "higher" or "same" and an unfavorable response was "lower." H1 was an "H" scale (For a discussion of the "H" technique, cf. 15.) with 4 triplets, no picture being used more than once. H2 was an "H" scale with 9 triplets, using 19 pictures, 8 of which were used twice with different cutting points. Sample 1 (N = 67) was the half of the original sample on which the scales were constructed. Sample 2 (N = 72) was the replication half of the original sample. Rockford (N = 37) was a sample of students in a woman's college.

of the original sample with this picture excluded showed the original reproducibilities lowered only .004. The picture was kept in the analysis because it contributed to the ordering of symbols, although not the ordering of subjects. It should be noted that for selected pictures and particular combinations of cutting points scales with higher reproducibilities can be constructed from these data.

In summary, if one accepts the Guttman scale as a measurement model, perceptions of subjective status and of the status of material status symbols tend toward unidimensionality. The argument differs considerably from Veblen's mordant observations in 1899, but the conclusions are congruent:

The means of communication and mobility of the population now expose the individual to the observation of many persons who have no other means of judging his reputability than the display of goods . . . (17, p. 71)

Measurement of Subjective Status

While the previous paragraphs described our scales in terms of the prestige ordering of the photographs, Guttman scales also present an ordering of subjects. In our data we have an ordering of subjects from those who said that many pictures were above them to those who said that few were. Our subject order is a grouping in terms of subjective prestige.² The technical properties of this scale are the same, of course, as those discussed above, with one exception. Our replications were of the object ordering, not of the subject ordering. We have no data on subject-test retest reliability, while our different samples do provide replications for the picture order.

What does our scale measure? There is a case for face validity, if only because the subjects were told what dimension the investigator had in mind, a situation which does not hold in the usual scaling of attitudes. Nevertheless, for theoretical and common-sense reasons we would be pleased if the scale had some relation to other measures of social status, but we should not expect these relationships to be perfect, for we hope that we are measuring a particular aspect of social status, not status in general. Armed with the comfortable hypothesis that our instrument should show moderate positive relationships with other status measures, let us examine some correlates of subjective placement.³

² It should be noted that Hatt's unsuccessful attempt to scale analogous data (7) worked with a different type of subject order. His scales ordered subjects in terms of their tendency to accord high prestige to occupations *in general*. It is interesting to speculate on whether he might have found a scale if he had asked his subjects to judge the occupations as higher or lower in addition to asking them to rate them.

³ Subject scores were derived as follows: Scale positions on the two Guttman scales

TABLE 2
Associations with Subjective Status Placement

Variable	Chi Square	Phi
Interviewer's economic rating	31.85	.50
Interviewer's picture rating	28.26	.47
Husband's occupation	17.97	.37
Subject's age	11.36	.30
Subject's education	10.96	.29
Husband's education	10.09	.28
Subject's father's occupation	7.12	.23

Note: Interviewer's economic rating is a rating of the subject's socioeconomic status on the standard market research "ABCD" scale. Interviewer's picture rating is the scale position of the picture which the interviewer chose as most like the subject's living room. Husband's and father's occupations are rated on a modification of the North-Hatt prestige scale. See footnote 5. All chi square tests are significant at .01 level.

The scale seems to show moderate and significant associations both with standard objective correlates and with the special device (the "picture rating") designed as an objective counterpart to our subjective scale. It should be noted that none of the relationships is strong enough to warrant the use of these variables as substitutes or predictors for subjective status. Whether this is due to measurement error or the tendency to independence of subjective status perception is unknown.

When we hold our strongest relationship (the interviewer's rating of economic status) constant, and repeat the process, we find that: (a) the subject's age, (b) the picture rating, and (c) the status of the husband's occupation still show relationships which are significant at the .01 level. We may assume that although economic status is the strongest single relationship, some noneconomic component of the symbols (taste?) and the prestige of the husband's occupation also enter into these subjective perceptions. The most interesting survivor of the partial correlation process is age. As we shall later show that age has a very strong relationship with the ability to judge status, we tested the hypothesis that the higher status accorded themselves by older women was accounted for by the fact that they were poorer judges and tended to flatter themselves. However, the relationship between age and subjective status is still significant with a probability of less than .01 when both economic status and ability to judge

were converted into rank positions in the sample, and a subject was given as her subjective status position the median of her two ranks. This distribution was dichotomized and fourfold tables using dichotomized status measures were made. All chi square coefficients from fourfold tables reported in this article are corrected for continuity.

status are held constant. Why should this be so? Such a finding in China would merely be an affirmation of a cultural value, but in our society the status of an older person is equivocal. A possible interpretation is not that the old inflate their status, but that the young deflate theirs. In a society characterized by high upward mobility and even higher orientations toward mobility, it might be that the young, having their eyes fixed on a stratum to which they aspire, are more conscious of the gap between their present position and positions above it.

As for the purely "status" variables, we note that it is the husband's achieved status which remains significant, while the spouses' ascribed statuses (their parents' occupations and her education, which presumably terminated before her marriage) "disappear" when economic status is held constant. Thus, it would seem that the wife's perception of her husband's occupational success is an intervening variable. This reinforces the statements of such writers as Parsons (13) and Davis (4), who have stressed that the status of the nuclear family is assigned to all on the basis of the occupational position of one, the husband.

Differential Ability to Judge Prestige

Having examined our subjects' self-judgments, let us turn to the question of their judgments of others. Now, while the existence of scalability assures us that the photographs were not piled at random, it is important to note that high scalability is compatible with considerable variation in perceptions of the total prestige continuum. A Guttman scale is concerned only with a dichotomous distinction, and is quite oblivious to how a subject might order items *within* each of the two groups. At an extreme, a subject who said she was higher than all the pictures could be a "perfect scale type" and still rank the pictures in an order which had a -1.00 rank correlation with the scale order. A subject with a median scale position could have a rank correlation in the neighborhood of $.50$ and still be a "perfect scale type."

Since our data consisted of sortings into four status groups as well as "higher" and "lower," we computed correlations between the general scale order of the pictures and a given subject's groupings.⁴ We will call these correlations measures of the ability to judge status, assuming that women whose correlations are higher (that is, whose card sortings were more like the general scale order) have a more accurate perception of the social prestige system than do those with lower correlations. The median coefficient was $.75$, a value which accounts for 56 per cent of the variance.

⁴ Product-moment correlations were computed by assigning the arbitrary values of 0 through 3 to the four piles. The rank order of the pictures was derived from the two Guttman scales by means of the operations described in footnote 3.

TABLE 3

Mean Correlation with the Prestige Order of Pictures by Age and Status Group

Status	Age		
	21-35	36-51	52-65
High	.803	.743	.733
Middle	.796	.729	.525
Low	.789	.647	.466

Therefore, for most of the women, most of their variance is accounted for by picture variation in the original scale.

We are studying perceptions of social structure. In order to analyze variations in these perceptions, we will examine certain structural variables. The American social structure seems to be organized on two axes, stratification and kinship role, and we shall look at both types of variables. Furthermore, we shall divide our structural variables into two types—those which measure location in the social structure (status and “domesticity”) and those which measure change in structural location (age, “vertical” mobility, and “horizontal” mobility). Our assumption is that both location and change in location are associated with differences in status perception.

Our sample was designed to test two hypotheses, one about location and one about change in location. The hypotheses were (a) the ability to judge status increases with status, and (b) the ability to judge status decreases with age. The three age groups were 21-35, 36-51, and 52-65. Status was measured by the North-Hatt prestige ranking of the subject's husband's occupation. Three arbitrary occupational strata were used.⁵ The “high” group consisted primarily of professionals and prosperous businessmen, the “middle” group consisted primarily of white-collar and skilled workers, and the “low” group consisted primarily of semiskilled and service workers.

Our design gave us 15 subjects in each of the 9 cells. The 5 refusals and 1 protocol which was unusable because of clerical mistakes left us with 11 cases in the low-status older age cell, and 14 in the middle- and high-status older age cells.

Table 3 shows the mean correlation with the scale order of the pictures for the nine groups.

The orderings of the cells confirm both predictions, but we should like to make a statistical test of significance for the differences through analysis

⁵ To be specific, the occupational ratings were taken from a Boston area replication of the North-Hatt instrument, cf., (1). Campbell's occupational ranking was divided into three groups, and occupations not included in his data were assigned by the writer by means of a typology derived from his findings.

of variance. Because the means and variances of correlation coefficients are correlated, in order to meet the assumptions of the technique, we transformed our data using Fisher's z' statistic. Bartlett's test of homogeneity of variances showed that the transformed-within-cell variances were significantly different ($p < .05$), and inspection of the table indicated that these variances were still negatively correlated with the means. We may infer that within our groups of poorer subjects, the women are not only less efficient judges but show greater individual variation.

Actual computation of analysis of variance did show that the prediction about age was significant at the .01 level, the prediction about status was significant at the .05 level, and interaction was insignificant, but we shall rest our case on a nonparametric test of the following hypotheses:

1. Subjects of the same rank standing within their cell will have higher correlations in younger groups than in older groups of the same status.
2. Individuals of the same rank standing within their cell will have higher correlations in higher status groups than in lower status groups of the same age.

For example, these hypotheses predict that the subject with the third highest correlation in the "young, high-status" group will have a higher correlation than the third highest subject in the middle-age, high-status cells the younger middle-status cell; and the younger, low-status cell. The null hypothesis is that one half of these predictions will be correct, and significance is computed from tables for the "sign test" (12). Table 4 presents the results.

We may consider the hypotheses to be substantially, if tortuously, demonstrated, although we would prefer that the status hypothesis be qualified by noting that it is the difference between the high-status group and the other two which is probably crucial. Despite the general tendency of the sample to perceive a status order of prestige symbols, individual

TABLE 4
Nonparametric Tests of Age and Status Hypotheses

Prediction	Number of Predictions	Number Incorrect	One-tailed Probability
All status predictions	126	36	< .01
High status vs. middle	44	8	< .01
High status vs. low	41	8	< .01
Middle status vs. low	41	20	> .125
All age predictions	123	11	< .01
Younger vs. middle	45	2	< .01
Younger vs. older	39	2	< .01
Middle vs. older	39	7	< .01

variations in perception are appreciable and are associated with the subject's social structural position and experience.

When age and status were held constant, certain other social attributes still have a significant relationship with the ability to judge prestige symbols. We shall discuss four, three of which are measures of change in location and one of location itself.

Our finding, that with status held constant the ability to judge status seems to decrease with age, raises many problems about the effect of change in structural position on perceptions of status. Two contradictory hypotheses about this relationship can be justified from the literature. On the one hand, one may consider the proposition that change in location is associated with greater ability to judge status. Sorokin, for instance, considers increased knowledge as an *effect* of social mobility (14, p. 509). The Warner "school" argues that heightened perception of status symbols may be a *cause* of change in location (16, p. 74). On the other hand, one may consider the proposition that change in location is associated with a lesser ability to judge status. Our age data may be interpreted this way. Hollingshead, for instance, excluded mobile judges from his Elmtown study, presumably because they were poorer judges (9, p. 33). Greenblum and Pearlin present data in their study of prejudice which would be congruent with this second hypothesis. They conclude:

It was found that mobility, up or down, results in an increase of prejudice. This increase was seen as a result of the development of prestige insecurity in the process of mobility. (6, p. 491)

We have data on three aspects of locational change to test these two hypotheses. In order to hold age and status constant, we ranked the subjects in each cell of our design and divided them into "better" and "poorer" judges by splitting them at the cell medians.

Let us first look at occupational mobility. We defined mobility as a difference between current prestige and prestige of family of origin, using our high-middle-low classification of occupational prestige. When we compare the subject's father's occupation with her husband's current occupation, we find that mobility, either upward or downward, is associated with poorer ability to judge status, but the difference is not statistically significant ($p < .10$). We do, however, get a clear-cut pattern when we taken into consideration, not the wife's mobility alone, but the mobility pattern of the couple. In Table 5, a husband's mobility was measured by comparing his occupation with his father's. A wife's mobility was measured by comparing her husband's occupation with her father's.

Apparently it is the relative mobility of the pair which leads to poorer prestige judgments for the mobile wife, although we do note a majority of "poorer" judges in the group where both were mobile in the same direction.

TABLE 5

Spouses' Relative Mobility and Wife's Ability to Judge Prestige (Age and Status Held Constant)

Relative Mobility	Better Judges	Poorer Judges
Both stable	33	16
Both mobile in same direction	15	19
One stable, one mobile	11	22

Note: Chi square = 14.455. $p < .01$.

We shall refrain from interpretation until we have presented data on "horizontal" mobility. We collected a number of measures of residential mobility (number of houses ever lived in, number of communities, etc.). When these were run against ability to judge status (with age and status held constant) we found no association. But we did find that intent to move, that is, the subject's expectation of future horizontal mobility, has a significant association with the ability to judge status ($p < .05$). Within each age and status group, women who expected to move in the near future were better judges of status. Apparently it is current orientation toward movement which makes the difference, not the amount of past shuttling through space.

Our four measures of change in location in the social structure seem to have no simple relation to the ability to judge social status. If considered as indices of distance traveled, they have contradictory implications. We suggest that one way of arranging them would be as follows:

1. The relationships with age, intent to move, and lack of relationship with housing history argue that perception of the symbol status order is more a function of current expectation of change in structural location than of past experience of such change. Thus, it may be that it is not that the older woman's past has lowered her correlations, but that she has less expectation of a social "future."

2. The data on spouses' mobility, coupled with the inference that the husband's occupation was an intervening variable in perception of subjective status, argue that: the nuclear family is the crucial unit of status and status perception, and movement in the social structure which creates differential perspectives within the nuclear family tends to lower a wife's ability to judge social status.

Having dealt at some length with location in the status hierarchy and with various changes in position, let us turn briefly to another measure of location. We assumed that prestige and kinship were the main axes of the American social structure, but our sample of married adult women places severe limits on exploration of kinship role as a variable. Nevertheless, we

TABLE 6
Child-Years and Ability to Judge Prestige (Age and Status Held Constant)

	Better Judges	Poorer Judges
More child-years	16	38
Fewer child-years	44	27

Note: Chi square = 11.60. $p < .01$.

do have information on one structural variation within the role of female spouse. Several measures—wife's employment, number of children, family size, length of marriage, etc.—suggested that wives who have spent more of their time with their children and less in the "outside world" are poorer judges. These considerations suggested as a general index of "domesticity" *child-years*—the sum of the ages of a subject's children. Women with higher sums, with age held constant, have spent more of their lives taking care of children, whether they reared more children or fewer children longer. Each age and status group was divided into women with more child-years and women with fewer. Table 6 results. It would seem that within an age and status group, the ability to judge prestige symbols is inversely related to domesticity.

SUMMARY

This study, which was based on the judgments of photographs by 134 housewives, attempted to demonstrate that:

1. There is a tendency for judgments of status symbol prestige to form a unidimensional order.
2. A picture technique may be used to measure one aspect of social status, subjective status perception.
3. Location in and movement within the social structure are associated with individual differences in the ability to perceive this unidimensional order. Indices of location are social status and domesticity. Indices of movement are age, experienced vertical mobility, experienced horizontal mobility, and expected horizontal mobility.

Manuscript received: January 31, 1956

Revised manuscript received: April 13, 1956

James A. Davis
Community Health Project
Harvard School of Public Health
20 Whittier Street
Boston 20, Massachusetts

REFERENCES

1. Campbell, J. D., "Subjective Aspects of Occupational Status," Ph.D. dissertation, Harvard University, 1952.

2. Centers, R., *The Psychology of Social Classes*, Princeton: Princeton University Press, 1949.
3. Chapin, F. S., *Contemporary American Institutions*, New York: Harper and Brothers, 1935.
4. Davis, K., *Human Society*, New York: The Macmillan Company, 1949.
5. Gross, N., "Social Class Identification in the Urban Community," *American Sociological Review*, 1953, 18, 483-497.
6. Greenblum, J., and L. Pearlin, "Vertical Mobility and Prejudice: A Socio-Psychological Analysis," in R. Bendix and S. Lipset (eds.), *Class, Status, and Power*, Glencoe, Ill.: The Free Press, 1953.
7. Hatt, P., "Occupation and Social Stratification," *American Journal of Sociology*, 1950, LV, 533-543.
8. Hetzler, S., "An Investigation of the Distinctiveness of Social Classes," *American Sociological Review*, 1953, 18, 493-497.
9. Hollingshead, A. B., *Elmtown's Youth*, New York: John Wiley, 1949.
10. Kahl, J. A., and J. A. Davis, "A Comparison of Indexes of Socio-Economic Status," *American Sociological Review*, 1955, 20, 317-325.
11. Lynes, R., *The Tastemakers*, New York: Harper and Brothers, 1954.
12. Mosteller, F., and R. Bush, "Selected Quantitative Techniques," in G. Lindzey (ed.), *Handbook of Social Psychology* (Volume 1), Cambridge, Mass.: Addison-Wesley Publishing Company, 1954.
13. Parsons, T., *Essays in Sociological Theory Pure and Applied*, Glencoe, Ill.: The Free Press, 1949.
14. Sorokin, P., *Social Mobility*, New York: Harper and Brothers, 1927.
15. Stouffer, S. A., E. F. Borgatta, D. G. Hays, and A. F. Henry, "A Technique for Improving Cumulative Scales," *Public Opinion Quarterly*, 1952, 16, 273-291.
16. Warner, W., and associates, *Democracy in Jonesville*, New York: Harper and Brothers, 1949.
17. Veblen, T., *The Theory of the Leisure Class* (New York: The Macmillan Company, 1899). Mentor edition, New York: The New American Library of World Literature, 1953.

Patterns of Perceived Interpersonal Relations¹

ROSALIND DYMOND CARTWRIGHT, *University of Chicago*

JULIUS SEEMAN, *George Peabody College for Teachers*

DONALD L. GRUMMON, *Michigan State College*

Many current personality theories take the position that the self is a product of social interaction, composed of self-other patterns and their intrapersonal organization. The study to be reported here was conceived to develop ways of isolating and characterizing these internalized patterns of expectation and response and to investigate the effects on them of psychotherapy experience. Such a study requires the spelling out of the dimensions along which these patterns might be objectively measurable and a method of obtaining a sampling of them which maximizes the possibility of getting comparable samples at different times and from different people. The method chosen was the Thematic Apperception Test (TAT). Since this test offers a series of standard social and nonsocial stimulus situations which each subject organizes in his story-response in his own way it provides a rich source of projected interpersonal relations from which many variables can be abstracted for intensive study.

Some of the dimensions along which these patterns were expected to vary were: relationship frequency potential (the number of times the people projected into the stories were actually in a relationship in contradistinction to acting without reference to others); the degree of differentiation of relationships (the variety of roles within which the persons in the relationships acted); activity-passivity (which of the participants, the hero or the other, initiated the action involved in the relationship); whether the activity appeared to be successful or unsuccessful for each of the participants; the nature of the affect of the characters; the strength of the affect; the manner in which the affect was expressed (constrained, explosive, or easily flowing); the dynamics of the actors, dependency, dominance, avoidance, aggression, etc.; the amount of social-emotional distance between the characters; the degree of satisfaction of each participant; the

¹ This study was conceived, the instruments built, and the data gathered under a grant from The Rockefeller Foundation, through its Medical Division, to the Counseling Center of the University of Chicago. The analysis of the data was done under a research grant (PHS M 903) from the National Institute of Mental Health of the National Institutes of Health, Public Health Service.

Thanks are due to the following people who helped in various stages of this study: Melvin Feffer, Leonard Hersher, Samuel Sutton, Robert Bauer, John Meany, Thomas Lynch, and Alice Wagstaff. Dr. Desmond S. Cartwright acted as consultant on statistical matters.

way in which the characters were described, in terms of their behavior, their feelings, or both. To investigate these various aspects a series of rating scales was devised which could be applied to each relationship found in each TAT story. (Copies of this rating form can be obtained by writing to the first listed author.)

THE SAMPLE

The sample for this study consisted of 21 persons who were self-referred to the University of Chicago Counseling Center for psychotherapy (5). There were 10 males and 11 females ranging in age from twenty-one to forty, the mean age being twenty-seven. Twelve of the subjects were students, the remainder having a variety of other occupations. The average number of interviews to complete therapy was 28, the range being 6 to 64. When the counselor involved in each case rated the success of the therapy on a 9-point scale (where 1 = complete failure and 9 = marked success), the range of success represented in this group was 2 to 9 and the mean was 6.

The subjects had all taken a number of tests, both before beginning counseling and at a follow-up point six months to one year after the completion of their interviews; among these was a standard 20-card TAT.

METHOD

The unit of analysis was a relationship. Each relationship involved two ratings on each scale for which there was evidence in the story, one for the hero and one for the other person involved. Taking the before-therapy and follow-up records of the 21 clients there were 40 stories available for rating from each subject, making a total sample of 840 stories. Each story, however, contained a varying number of relationships, the range being from 0 to 5. Each story was coded individually and separated from the others in the record. In this way the ratings of the relationships could be made without the judges having any knowledge of whether the particular story was from a before- or after-therapy record and without any halo effect from the other stories which were told on the same occasion.

RELIABILITY

Five judges made all the ratings of the more than 2000 relationships involved. After an initial training period in the use of the scales their interjudge reliability was tested by having each of them rate one relationship in each of the same 15 stories. Since each relationship involved 2 individuals there was a total of 30 judgments per scale on which the extent of their agreement or disagreement could be computed. Cartwright's alpha (2), a nonparametric interjudge reliability measure, was computed for each

scale. This involves computing the actual number of agreements between judges and dividing by the maximum number of agreements possible. In this case, with five judges there are ten possible agreements per item. Three scales were found to be used with insufficient reliability to be worth retaining. The alpha for the remaining scales ranged from .42 to .69 for which the levels of significance ranged from .07 to .005.

ANALYSIS

The analysis involved a search for clusters of interrelationships in the variables studied. Two questions were investigated: (a) Are there distinct patterns in the way interpersonal relations are viewed by a group of clients after they have had psychotherapy and a chance to consolidate this experience and (b) are there distinct patterns in the way these perceptions change over a time interval which includes psychotherapy and so, presumably, a period of considerable change? To answer these questions the method developed by Butler known as rank pattern analysis (1) was used. This method was designed to isolate clusters in data where the basic units for quantification are frequencies. In the reference cited the close similarity of rank pattern analysis to multiple factor analysis is shown by applying the former to data which have already been factored by the latter method.

Two rank pattern analyses were done: (a) of the follow-up status of the 21 subjects on all the reliably rated aspects of their projected interpersonal relations, and (b) of the change between the pretherapy and follow-up status of these people on these dimensions. Since Butler's paper describing this method has not yet appeared in published form the steps involved in pattern extraction as they applied in this study are briefly outlined in Appendix A.

Table 1 gives the pattern structure of the follow-up matrix. For quick interpretation of the aspects which varied together, the direction in which each variable was ranked is given in brackets. In all cases the rankings were made with rank 1 representing the favorable and rank 21 the unfavorable extreme of the variable. Table 2 gives the structure of the matrix of change from pretherapy to follow-up.

Following the extraction of the patterns the next step in the procedure was to assign a score to each subject on each pattern. This involved converting the original ratings into T scores so that they might be summed. The T scores for those ratings which formed a pattern were then summed for each subject and the 21 subjects then ranked on their pattern scores.

RESULTS I

From Tables 1 and 2 it appears that the answer to both questions is clearly affirmative. There are distinct, orthogonal patterns of perceived

interpersonal relations. (The intercorrelation between the subject's pattern scores is zero.) The question of the statistical significance of these patterns is a difficult one as there is no generally stable solution available in factor analysis (7). The most commonly used criterion is that of Guilford and Lacey (3) which requires that the product of the two highest loadings on a factor exceed the standard error of zero r . This criterion is applicable to multiple group factors, the method for which closely approximates the rank pattern method (6). Since Spearman's rho is a close rank approximation to r , and the distribution of rho approaches normality with $N \geq 20$ (4, pp. 46-69), it is plausible to estimate the significance of these patterns by a criterion analogous to that of Guilford and Lacey.

Since the standard error of zero rho is $\sigma p = \sqrt{1/N - 1}$ for a pattern to be significant the product of its two highest loadings must be greater than $\sqrt{1/N - 1}$. In the present case with $N = 21$, $\sigma p = .224$. However, the criterion for the inclusion of a variable in a pattern was that its correlation with the average rank of all other variables in the pattern be not less than .498 (see Appendix A). Thus, even the product of two such minimum loadings would exceed the figure required to fulfill the Guilford-Lacey criterion (.498² = .246). It seems therefore that the patterns as isolated are significant.

To determine just how significant each pattern may be requires a different technique. In this case Kendall's W -coefficient (4, pp. 80ff) was used. Rank patterns are obtained from the correlations between variables ranked in order of size of the correlation with all other variables in the matrix. For any matrix such row rankings are invariant over all permutations of the order of variables in the matrix. This means that the column sums are also invariant. Treating rows as observers (m) and variables or columns as entities ranked (n), W is computed for each pattern by the

formula $W = \frac{12S}{m^2(n^3 - n)}$ and the significance of the W coefficient obtained tested by the formula $\chi^2 = m(n - 1)W$ with $n - 1$ degrees of freedom. Table 3 summarizes these results.

Having established that these patterns are significant let us look first at the three patterns which emerged from the rank pattern analysis of the follow-up TAT's. Pattern A appears to be one of external description of relations. Pattern B seems to involve affective release. Pattern C seems best summarized as a comfortable adjustment pattern. In pattern A the following aspects were found to vary together: a high potentiality for seeing people to be in relationship to one another, a high differentiation of their roles, a high degree of communication between the characters, a high differentiation of the dynamics involved on the part of both participants, and a relative absence of feeling descriptions. This pattern appears to be one of sensitivity to the perception of relationships, but this perception is

TABLE 1
*Pattern Structure of the Follow-up Matrix**

Variables	Pattern A	Pattern B	Pattern C
1. Total no. of relationships in the follow-up TAT (high)	.856*	-.113	-.191
2. Amount of communication present in projected relationships (high)	.679*	-.059	-.222
3. Number of different kinds of relations (high)	.790*	-.033	-.350
4. Number of different dynamics used to describe heroes (high)	.834*	-.036	-.126
5. Number of different dynamics used for others (high)	.819*	.228	-.625
6. Number of descriptions of heroes in feeling terms (low)	.580*	-.712	.424
7. Number of descriptions of others in feeling terms (low)	.849*	-.164	-.259
8. Number of "no evidence" ratings on all scales (low)	.119	.913*	-.822
9. The hero as the source of activity in relationship (high)	.325	.670*	-.365
10. Number of different affect states descriptive of hero (high)	-.358	.918*	-.632
11. Number of different affect states descriptive of others (high)	-.217	.913*	-.748
12. Use of positive dynamics as descriptive of others (high)	-.317	.563*	-.248
13. Amount of affect expressed as easy flowing (high)	-.399	.857*	-.546
14. Number of relations of hero rated as satisfactory (high)	-.178	.625*	-.181
15. Description of heroes in behavior terms (low)	-.222	.896*	-.715
16. Description of heroes in both behavior and feeling terms (high)	.085	.887*	-.709
17. Description of others in both behavior and feeling terms (high)	.232	.794*	-.792
18. The ratio of positive to negative affect (high)	-.079	-.772	.925*
19. Use of negative dynamics as descriptive of hero (low)	-.399	-.591	.797*
20. Use of negative dynamics as descriptive of others (low)	-.742	-.161	.624*
21. Use of positive dynamics as descriptive of heroes (high)	.198	-.379	.505*
22. Amount of constrained expressivity of affect (low)	-.599	-.196	.720*

* *Note:* All members of a pattern are starred in the column headed by the appropriate letter. It should be noted that negative loadings are interpreted as zero loadings in rank pattern analysis.

TABLE 1—*Continued*

Variables	Pattern A	Pattern B	Pattern C
23. Amount of explosive expressivity of affect (low)	-.064	-.657	.498*
24. Amount of extremely weak and/or strong involvement of heroes (low)	-.212	-.724	.869*
25. Amount of moderate involvement of heroes (high)	.425	-.601	.593*
26. Number of relations unsatisfactory to hero (low)	-.201	-.583	.771*

of what can be seen from the outside without much personal involvement. It also appears to say very little about the quality of the relations, being concerned primarily with quantitative aspects.

Pattern B is very different in that it focuses on affectivity both as it is perceived and expressed. In this cluster the elements which vary together are: a high degree of self-motivation of the heroes; the relations giving satisfaction to the heroes; affect being highly discriminated and expressed easily; expected positive responses from others, sharing, approaching, giving, enjoying, etc.; and including both feeling and behavior in describing the interaction of the participants. It is interesting to note that this is a cluster which might be expected to be related to client-centered therapy with its emphasis on the client's ability to manage his own therapy (self-motivation), if he becomes aware of his feelings (differentiation of affect), and is able to express them freely (easy flow of affect), and so integrate them into his awareness of self and others (behavior and feelings descriptions).

Pattern C seems to describe a state of comfort and good will of heroes to others and the absence of maladjustive aspects in the relations. It asserts that the affect in the perceived relations is more positive than negative, that the self-figures respond to others in positive dynamic terms and that their attachments to others are moderately strong, that the negative responses—aggression, dependence, rejection, rebellion, etc.—are relatively little involved, that constrained or explosive expression of affect is relatively absent as are relations unsatisfactory to the heroes and involvement of the heroes which is overly weak or excessively strong. This pattern does not define the positive adjustment as clearly as it does the poor adjustment which is less present in those who rank high on this pattern.

There is considerable similarity between the three patterns described above and the four which were derived from the change matrix. Pattern D

TABLE 2
Rank Pattern Structure of Pretherapy to Follow-up Change Matrix

Variables	Pattern D	Pattern E	Pattern F	Pattern G
1. Increase in number of relationships	.788*	.038	-.493	.077
2. Increase in number of kinds of relationships	.721*	-.595	-.107	.442
3. Increase in ratio of positive to negative affect	.617*	.211	-.400	.166
4. Increase in number of different dynamics descriptive of heroes	.713*	.153	-.232	-.098
5. Increase in number of different dynamics descriptive of others	.729*	.395	-.428	-.159
6. Decrease in descriptions of others in feeling terms	.637*	-.101	-.363	.319
7. Decrease in frequency of "no evidence" ratings	.165	.883*	-.777	-.741
8. Increase in number of affect states descriptive of heroes	-.401	.780*	-.195	-.598
9. Increase in number of affect states descriptive of others	.156	.622*	-.695	-.359
10. Increase in amount of affect expressed as easy flowing	-.310	.826*	-.382	-.707
11. Increase in amount of moderate involvement of heroes	.334	.665*	-.428	-.529
12. Increase in number of relationships satisfactory to heroes	.034	.888*	-.551	-.787
13. Increase in number of relationships satisfactory to others	-.084	.929*	-.432	-.782
14. Decrease in number of descriptions of heroes in behavior terms	-.168	.895*	-.472	-.751
15. Increase in number of descriptions of others in both behavior and feeling terms	.337	.765*	-.769	-.406
16. Increase in number of times the hero is source of activity	-.136	-.615	.573*	.338
17. Decrease in number of descriptions of others in negative dynamics	-.555	-.373	.773*	-.016
18. Decrease in the amount of constrained expressivity of affect	-.360	-.598	.913*	.415
19. Decrease in the amount of extremely strong and/or extremely weak involvement of heroes	-.362	-.250	.701*	.104
20. Decrease in number of relations unsatisfactory to hero	-.321	-.695	.842*	.515
21. Decrease in number of relations unsatisfactory to others	-.433	-.675	.764*	.496
22. Increase in amount of communication present	.169	-.758	.247	.794*
23. Decrease in description of heroes in negative dynamics	-.185	-.761	.481	.549*
24. Decrease in description of heroes in feeling terms	.441	-.644	.204	.707*

TABLE 3
The Significance of the Patterns

Pattern	<i>W</i>	χ^2
A	.707	133.623
B	.715	193.050
C	.575	139.766
D	.619	89.136
E	.770	166.358
F	.709	102.124
G	.738	53.164

Note: The p value for all patterns is less than .001.

is very like pattern A in that five of the six variables of which it is composed are also in pattern A. It, too, can be characterized as an external descriptive pattern. Pattern E shares seven of its nine variables with pattern B and focuses on increased affectivity, and pattern F shares four of its six variables with pattern C. It seems, like C, to be an adjustment pattern. Pattern G is mixed, sharing two of its variables with A and one with C. This tentatively can be called a decreased negative response pattern.

Pattern D shows that the potentiality for seeing relationships, the differentiation of roles and dynamics in both heroes and others, changes in the affect ratio, and the potentiality for seeing others in pure feeling terms are aspects which all change together. This pattern, like A, except for the more positive feeling change, is one of seeing people externally.

Pattern E includes increased differentiation of affect and its easy expression, increased satisfaction in relations, and moderate involvement in them.

Pattern F, like pattern C, seems to involve change in the direction of projecting less maladjustment and discomfort in relations. The related changes here are: less expectation of negative responses from others, fewer relations that are unsatisfactory to both participants, with constraint on the expression of affect, and with the social-emotional distance of the heroes going to extremes of very involved and very uninvolved. As these things changed for the better (became less prevalent in the stories over therapy), the heroes became more self-motivating. This would seem to be a pattern that involves freeing the subject from his own passivity which has been unsatisfactory to him (constraint, self a minor source of activity and at the mercy of others motivated by negative dynamics).

Pattern G looks like a straight description of therapy behavior: as communication increases, one's description of self-figures becomes less loaded with pure feeling and less often are self-figures rejecting, rebelling, attacking, dominated, or domineering in their relations to others.

RESULTS II

Having found that there are distinct independent significant patterns of ways of perceiving interpersonal relations and patterns in the way these change, the relation of the pattern scores to other variables was investigated.

The relation of the patterns scores to the age, sex, and occupational status of the clients was found to be insignificant. Length of therapy did show some relationship to patterns F and G. Tables 4 and 5 show that when the pattern scores are dichotomized at the median and the cases split into long and short therapy a chi-square test is significant at the 5 per cent level ($\phi = .42$). The longer cases show more change of the kind represented by patterns F and G.

A further interesting finding in relation to pattern F came to light when the therapist of each case was listed. In this group of 21 cases, 10 different therapists were involved. However, three of them (X, Y and Z), had three, four, and five cases respectively, and the remaining nine cases were divided among seven counselors. The three cases of counselor X all ranked very high on their pattern F scores: 1, 2, and 4. Three of the cases of counselor Y ranked in a group in the middle range: 12, 13, and 14. Three of the cases of the third therapist Z ranked at the bottom: 19, 20, and 21. It appears that although all therapists used the same technique, client-centered therapy, different therapists contribute differently to the kind of changes represented by pattern F.

TABLE 4
Relation of Pattern F to Length of Therapy

Pattern F	Length of Therapy	
	Short (Less than 27 Interviews)	Long (More than 27 Interviews)
High scores (1-10)	3	7
Low scores (11-21)	8	3

$$\chi^2 = 3.83$$

TABLE 5
Relation of Pattern G to Length of Therapy

Pattern G	Length of Therapy	
	Short (Less than 27 Interviews)	Long (More than 27 Interviews)
High scores (1-10)	3	7
Low scores (11-21)	8	3

$$\chi^2 = 3.83$$

That this is not an accidental finding resulting from counselor X's cases all happening to be very successful in therapy and counselor Z's very unsuccessful is demonstrated by the therapists' own rating of the cases on the 9-point scale of success. X rated his three cases 8, 6, 6 (mean = 6.6), while Z rated his three 7, 7, 7 (mean = 7).

Since this finding suggested that the three different counselors might be having a differential impact on the way in which their own clients changed their patterns of perceived interpersonal relations, a method was sought which would clarify this process. Kendall's paired comparison method was used and his coefficient of agreement (4, pp. 125-131) computed. In this case the ranks of change on the four change patterns were compared for the clients of a particular therapist to determine if one type of change were "preferred" over another by examining all possible pairs. For example, counselor X saw three people—how many of them changed more (had a rank closer to 1) on D than on E, on D than on F, on D than on G, etc.?

The results of this analysis were that counselor X's clients preferred the change patterns in the order F, G, D, E, and this was significant at the .03 level. Counselor Y's four clients preferred the patterns in the order D, E, G, F, a result which was also significant at the .03 level. The preferences of counselor Z's clients were nonsignificant. Without doubt, counselor X's clients change most in the direction of comfortable adjustment to others, less on decreased negative response to others, and least on external description and affective release. Counselor Y's cases, on the other hand, change most on external description and affective release and less on decreased negative response and comfortable adjustment. There is no pattern to the changes of counselor Z's cases as a group. The full significance of this result becomes more striking when it is remembered that the time interval involved in the change scores not only included the time during which the subject was in therapy but also a period of six months to one year following this during which there was no contact with the therapist. This means that the effect of therapist X and Y in changing the way their clients' perceived people interrelating was not only consistent but persistent for a fairly long period. Having found that there are differential amounts of impact of the counselor and impact of different kinds it is now possible to study the transcribed interview material to ferret out the differences in therapist behavior that produces these results.

SUMMARY AND DISCUSSION

This has been a study of the patterns of expectation and response in the perceived interpersonal relations of 21 people who have completed a course of psychotherapy and a follow-up period and the patterns of change

in these perceptions over this period. It has asked two questions: Are there distinct ways of viewing interpersonal relations which can be isolated, and, are there distinct patterns in the way these perceptions change?

Seven patterns were isolated, three independent significant patterns of follow-up status and four of change from pretherapy to follow-up. The three follow-up patterns were called external description, affective release, and comfortable adjustment and the four change patterns, external description, affective release, comfortable adjustment, and decreased negative response. When the 21 subjects were scored on these patterns it was found that those who had longer therapy increased more on the comfortable adjustment pattern than those whose therapy was short. Also the longer cases decreased more on the negative response pattern than did the shorter cases. Further, it appeared that therapists differ in the extent and kind of their impact on the perceptions of self and others of their own clients.

Manuscript received: January 6, 1956

Revised manuscript received: March 28, 1956

Rosalind D. Cartwright
Counseling Center
University of Chicago
5737 Drexel Avenue
Chicago 37, Illinois

REFERENCES

1. Butler, J. M., "The Rank Pattern Analysis of Therapeutic Protocols" (dittoed paper), Counseling Center Research Files, University of Chicago, 1951.
2. Cartwright, D. S., "A Rapid Non-parametric Estimate of Multi-judge Reliability," *Psychometrika*, 1956 (in press).
3. Guilford, J. P., and J. T. Lacey (eds.), "Printed Classification Tests," *Army Air Forces Aviation Psychol. Program*, Res. Rep. No. 5, Washington, D. C.: U. S. Government Printing Office, 1947.
4. Kendall, M. G., *Rank Correlation Methods*, London: Griffin, 1948.
5. Rogers, C. R., and Rosalind F. Dymond (eds.), *Psychotherapy and Personality Change*, Chicago: University of Chicago Press, 1954.
6. Thurstone, L. L., *Multiple Factor Analysis*, Chicago: University of Chicago Press, 1947.
7. Vernon, P. H., *The Structure of Human Abilities*, London: Macmillan & Co., Ltd., 1951.

APPENDIX A

For the rank pattern analysis of the follow-up status, the 21 subjects were first ranked on each of the 35 scaled dimensions. These rankings were then intercorrelated. This procedure resulted in a 35 by 35 matrix of rho correlations. From this matrix seven rows were eliminated as not having sufficiently high intercorrelations with other variables to be worth retain-

ing. From the reduced 28 by 28 matrix three patterns were extracted by the following method:

1. The rows of intercorrelations were ranked according to size 1 to 28.
2. The rows which appeared by inspection to be highly similar were selected as a tentative pattern.
3. The average rank of this group of rows was computed by summing the ranks by columns and then ranking these sums.
4. The significant correlations of each row within the pattern was established by correlating each row of rankings with the average of the others within the pattern minus itself. These should be at the 5 per cent level of significance for acceptance of a row into a pattern. The actual cutting point used in this study was .498 which is slightly better than the 5 per cent level as given by Fisher for r .
5. Each of the other rows (not in the tentative pattern) was then correlated with the total average ranking of the pattern (see step 3). If all other rows have negative or nonsignificant positive correlations with the average and the condition set out in step 4 is met, then the tentative group of rows is taken to define the first pattern. It should be noted that negative loadings have the effect of zero loadings in rank pattern analysis.

This procedure was followed until all possible patterns were extracted.

The same procedure was followed in extracting the patterns from the second matrix which was made up of the intercorrelations of the ranked change scores from pretherapy to follow-up on all the reliably rated aspects of the relationships. This was reduced to a 25 by 25 matrix eliminating those rows where the correlations were negligible with the other variables. Four patterns were extracted.

Cognitive Similarity in Facilitating Communication¹

PHILIP J. RUNKEL, *University of Illinois*

It has long been commonly observed that the meaning of any spoken phrase or any gesture depends upon the expectations within which it is embedded. Another way to say this is that the stimuli comprised by a communication impinge upon a set of potential responses belonging to the listener. This set of potential responses, being a limited selection of all human acts and being organized by a hierarchy of probabilities, must mediate every communicative process by presenting a framework, or mechanism, within which any communicated message will find its effect. It is this framework of potentialities upon which are engraved expectations of culture and role and the demands of the situation. In the terms of this framework any communication finds its resultant.

Clearly, an act of communication is itself a response. The possibilities of response are the possibilities of communication. The transmitter as well as the receiver of communication acts within a limiting framework. Using the term "cognitive field" to designate the possibilities of response, we may conceive the total communication process as an interaction between cognitive fields. The stimuli which impinge upon each field bring about alterations of response not only to the stimuli explicit in the situation but also to stimuli which are carried implicitly in the field; and further, the response which we see as communication arises not only from the stimuli offered by the other communicator, but also from the many stimuli implicitly associated in the cognitive field of the speaker and from the hierarchy of potential responses which organize them.

It follows from this view that communication cannot fruitfully be conceived as a sequence in which self-contained packets of information are exchanged. It is not a process in which one person merely adds to the belongings of another by "giving" him information. It is rather a kind of guessing game. Each person carries with him his cognitive field as a map of the world. He responds not to the world, but to the map. When he receives the stimulus of a communication, the meaning it has is the way it can be fitted into the map. When the communication fits readily, one's confidence in his map of the world is increased.

Since the effects of a communication depend on the manner in which

¹ This paper reports some of the findings of a dissertation (5) submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Michigan. Dr. T. M. Newcomb, particularly, is herewith thanked for his unflagging encouragement.

it "meshes" with an existing cognitive map, we might entertain the notion that these effects will take place more readily when the cognitive maps of the communicators are similar in structure. In fact, the general hypothesis of this paper is that *similarity of structure between two cognitive fields increases the efficacy of communication between them*. In its general form, this hypothesis is no doubt as old as communication. The contribution of this paper to the problems of communication lies not in the general terms of the problem chosen for study, but rather in the forms by means of which quantification has been applied to similarity of cognitive structure. In the present study, the particular index which furnishes operations for assessing similarity of cognitive structure is one which I have labeled "co-linearity" and which will be explained below.

THEORY

The theoretical framework supporting this investigation of the mediation of communication is a deliberate attempt to utilize in one consistent scheme Newcomb's (4) theory of communicative acts and some of the ideas of C. H. Coombs, particularly as the latter are represented in the monograph by Coombs and Kao (3). Because of space limitations, only those theoretical concepts necessary for explaining the experimental operations will be presented here.

We define an *orientation* toward a stimulus A as the set of all potential responses which the individual might at some moment make toward stimulus A or a set of stimuli containing stimulus A. There may at any moment be a number of possible respects in which the individual might respond to stimulus A. That is, there may be a number of *attributes* of stimulus A and the other stimuli in the situation which are relevant at that moment. These *relevant attributes* serve as reference vectors and define a multidimensional space within which the individual's potential responses to stimulus A are determined.

Although the space in which an orientation occurs may be multidimensional, the individual may simplify the interrelations of his judgments or potential responses by combining in some way the attributes of the objects to which he responds. By weighting or ordering the relevant attributes in some way, the individual can frequently put a number of stimuli into a simple (or linear) order. When an orientation involving a set of stimuli is so highly structured that the individual responds to the stimuli as if they were simply ordered, an important consequence is that *certain responses to the stimuli will be impossible* to him as long as he maintains the cognitive structure within which the simple order is defined. This assertion is not so complicated as it may sound, and illustrations are easy to find. Most men, for example, perceive the "fit" of clothing to correspond to the sizes

in which the clothing is manufactured. Suppose a man is trying on suits of sizes 36, 38, 40, 42, and 44. If he decides that size 42 fits him best, it would be impossible for him to report that the size 40 suit was a *worse* fit than the size 36.

But now the reader may object. "Is it inconceivable," he may ask, "that some man might be found who would give such an unusual opinion?" Let us put ourselves in the shoes of such an unusual fellow. We would feel, I believe, that he had found some characteristic or attribute of the suits which made it reasonable for him most to prefer the size 42, next the size 36, and least the size 40. Whatever this attribute might be, it would certainly be above and beyond the characteristics determining the usual order of the sizes. In other words, the response of our deviant would enable us no longer to interpret the preferences of individuals within the unidimensional space indicated by the simple ordering of sizes, but would require us to admit a multidimensional response space. This illustration points up the two parts of the assertion made earlier that the individual will find certain responses impossible *as long as he maintains a cognitive structure which puts a simple (unidimensional) order on the stimuli*.

Now let us suppose that person B is communicating with person A about a set of stimuli. Suppose that a clothing salesman is communicating with a customer about suits. "Try on this size 40," the salesman suggests. "Not quite right? Well, let's try a size 42. There, that looks just fine, doesn't it? You'd like to try a size 44? All right, here you are. Yes, I agree, this one is a little too large. The size 42 is just right for you."

Let us now ask the customer what the salesman would say about a size 38 or a size 36, or about a size 46 or a size 48. Obviously, the customer can predict very well what sizes the salesman would consider too small and too large. The point is that in providing the customer with an attribute in terms of which his judgments are being made, the salesman is giving the customer his opinions about stimuli which are not mentioned explicitly. From a sample of observed stimuli, the customer *gets information about other stimuli which can be judged according to the same attribute* as that underlying the judgments among the sampled stimuli. The important qualification here is that both communicators must be making their judgments, and interpreting the communication which occurs, according to the same attribute. If this is the case, each person can make correct predictions about responses the other would make to stimuli not yet communicated about explicitly.

But the customer might be purchasing a suit to wear to a fancy-dress ball, and the order of sizes might not be at all the order in which he judges how funny the suits are. If the customer with such a purpose does not let the salesman know what attribute is underlying his judgments, we can

only feel sorry for the salesman when we imagine the communication which might take place.

When two communicating individuals utilize the same underlying attribute in forming their judgments, we shall say that their orientations are *co-linear*. When they utilize attributes which would give at least some conflicting judgments, we shall say that their orientations are *non-co-linear*. These terms are chosen for their metaphorical reference to the unidimensional requirement and serve to suggest that the individuals are "on the same line" or "not on the same line," respectively.

CO-LINEARITY

The concepts so far presented have brought us to the point of being precise about a way of investigating the proposition that similarity of structure between two cognitive fields increases the efficacy of communication between them. If, in a multidimensional space, two individuals select the same attributes as relevant, they may then (or they may not) compose or weight these attributes so as to resolve their judgments into a simple order. If they do this, thus resolving a multidimensional space into a unidimensional, they may or may not do so similarly. The resulting simple orders, in brief, may or may not be co-linear. But if the two individuals do form orientations which are co-linear, then communication from one to the other transmits information about stimuli *in addition* to those stimuli which are explicitly mentioned in the communication as was illustrated earlier. We can now state our central hypothesis more precisely. We can assert that *communication about a sample of stimuli will convey more information about the stimuli from which the sample is drawn if the orientations of the communicators are co-linear, and less information if they are non-co-linear.*

The index of co-linearity used in the present research is built on the unfolding technique of Coombs (1, 2). In terms of the illustration of the customer looking for a suit, the preferences of most men in regard to fit would "unfold" into the order in which sizes fall. A man on the small side of size 42 might prefer five suits in the order of the sizes 42, 40, 44, 38, 36. And a man near size 36 might have the preference order 36, 38, 40, 42, 44. But we can call out both these orders because we are aware of the underlying attribute which organizes the preferences. There are some permutations of these five numbers which cannot be called off by standing in the shoes of a man of some one size and calling off the sizes in order of best fit to worst.

Orders of five stimuli which can unfold with A B C D E into *no* underlying order are given in Table 1. This table is used to compare the rank orders yielded by the responses of two subjects. Every rank order in the

TABLE 1
Co-linearity Table For Five Stimuli

Given two rank orders, re-label the stimuli of one order A B C D E, respectively. Then label each stimulus of the second order with the letter assigned to that stimulus in the first order. If the resulting second rank order appears below, the two given rank orders are *non-co-linear*.

A B E D C	C E D B A
A C E D B	D A C B E
A D C B E	D A C E B
A D C E B	D A E C B
A D E C B	D B C A E
A E B D C	D B C E A
A E C B D	D B E C A
A E C D B	D E A C B
A E D B C	D E B C A
A E D C B	E A B D C
B A E D C	E A C B D
B C E D A	E A C D B
B D C A E	E A D B C
B D C E A	E A D C B
B D E C A	E B A D C
B E A D C	E B C A D
B E C A D	E B C D A
B E C D A	E B D A C
B E D A C	E B D C A
B E D C A	E C A D B
C A E D B	E C B D A
C B E D A	E C D A B
C E A D B	E C D B A
C E B D A	E D A C B
C E D A B	E D B C A

table is non-co-linear with the rank order A B C D E. For five stimuli, there are fifty rank orders which are non-co-linear with any given order, and seventy which are co-linear.

PROCEDURES

Students in the introductory course in psychology at the University of Michigan were presented with five statements which could be seen as related to the content of the course but which were not assertions of the kind that would be made as a part of the material to be learned in the course or given as items on tests. The five statements used are these:

1. The conditions of living in the United States tend to narrow the range of things we are able to decide to do, think about, etc.
2. People who have a firm moral code are in general better adjusted than those who have not.

3. The biggest weakness of present-day psychology is that it is too theoretical.

4. Individuals could be changed in practically any way one might wish if the environment could be appropriately controlled.

5. The strongest influence in shaping a person into the kind of person he becomes, is his mother.

For reasons which will appear shortly, the Method of Rank Order was not used in collecting responses to these statements. The method used was the Method of Triads (2, p. 502). The statements were presented in groups of three, all ten of the possible combinations being used. The subject was instructed to mark, in each triad, the statement with which he *most* agreed and the statement with which he *least* agreed. Data were collected in this way from the classes of five teachers during the first week of the semester, and the identical procedure was repeated during the last week but one of the semester. The same questionnaire given to the students was also given to each of the five teachers. The preferences of teacher and student among the five statements were tallied, and a rank order inferred in each case where there was no evidence of a multidimensional response. Each teacher's rank order was then compared with the rank order given by each of his students by means of the co-linearity table, and the teacher-student pair was then categorized as co-linear or non-co-linear.

Since we have asserted that co-linear pairs of persons should exchange information more efficiently, and since quizzes on the course work can be taken as a measure of the degree to which the student has received the information which the teacher has given, we are ready to state hypotheses in operational terms. Since reliability or stability of a unidimensional orientation is one of the postulated requirements for a prediction based on co-linearity, two hypotheses will be advanced for testing, so that the results for each, when compared, will provide a check on the postulated effect of stability of the orientation.

Hypothesis 1: Among students who yield reliable rank orders of attitude items pertinent to the course, those who from pretest to posttest maintain rank orders co-linear with that of the instructor will receive higher grades on quizzes than those whose rank orders remain non-co-linear with that of the instructor.

Hypothesis 2: The difference in quiz grades predicted by Hypothesis 1 will be at least as pronounced when only those students are considered whose pretest and posttest rank orders are co-linear.

It is now appropriate to gather together a few loose threads; these will concern the selection of stimulus-statements for indexing co-linearity and the method of selecting stable and unidimensional responses from the responses obtained.

We have already mentioned that a direct sample of quiz material was

avoided in choosing the stimulus-statements for indexing co-linearity. Obviously, no elaborate theory would be required to predict quiz grades from a sample of quiz material. A further word, however, needs to be said about the way in which the stimulus-statements should be "related" to the content of the introductory psychology course. Briefly, the statements were chosen so as to "represent," in a special sense, the cognitive fields (or response spaces) which would mediate the communication between teacher and student. In order for a set of stimuli to be "representative," each statement should be multidimensional. That is, it should be possible for one subject to judge the statement on the basis of one attribute, and for another subject to judge the same statement on the basis of another attribute. With such stimuli, the order of preference given by the subject can reflect the attributes, and his weighting of them, which the subject brings to the stimulus situation.

To select a group of multidimensional statements, a long list of statements was first put together in which each statement, in the judgement of the experimenter, seemed interpretable from more than one viewpoint. These statements were then carried to a number of teachers of the introductory course in psychology, and each teacher was asked to state reasons which students might have for agreeing or disagreeing with the statements. The objective was to find a set of statements (a) which could be judged from a variety of reasons or viewpoints, and (b) could be discriminated from each other in regard to degree of agreement or disagreement with the statement. It will be seen that the search was for a highly heterogeneous set of items, rather than for a homogeneous set. The final selection rested on the judgments of the experimenter and the teachers. (A similar procedure was used in an earlier pilot study done with classes in zoology, which gave results substantially the same as those to be described here.) According to the theory, any set of stimulus-statements which was representative in the sense indicated, and discriminable, would have done as well as the set chosen.

We now turn to the method of selecting stable and unidimensional responses. The selection of subjects for whom co-linearity with the teacher was computed went through the following stages:

(1) Out of seven classes in introductory psychology (taught by five teachers), some students responded only to the pretest or only to the posttest. The number responding to both administrations of the questionnaire was 145.

(2) Of 145 subjects responding at both pretest and posttest, 15 gave responses at one time or the other which were intransitive, indicating that they were "unwilling" to compose the stimuli into a simple (or unidimensional) order. This left 130.

(3) The 130 transitive subjects gave responses which contained varying degrees of inconsistency. The Method of Triads presents each pair of stimuli to the subject three times when five stimuli are used. It is therefore possible for the subject to express a preference for stimulus A over stimulus B at one moment and for B over A at a later moment. If a subject is highly inconsistent, there is some ponderable possibility that the weight of his responses would have yielded an *intransitive* relation among the stimuli, had he responded a moment later than he did. In this sense, inconsistency may be interpreted as "uncertainty" on the part of the subject about putting a simple order on the stimuli. An arbitrary criterion was established at 70 per cent of the paired comparisons. Subjects who gave inconsistencies in 30 per cent or more of the pairs of stimuli were dropped from consideration. This removed 54 subjects, leaving 76. All five teachers gave transitive responses containing at least 80 per cent consistency.

(4) Hypothesis 1 makes explicit the next step in selection. Once the co-linearity index is applied to two rank orders, it provides in itself evidence for change of viewpoint between the two responses. (I shall frequently from here forward use the term "viewpoint" as a synonym for the cognitive structure underlying a rank ordering.) Subjects whose pretest responses were co-linear with the teacher's, but whose posttest responses were non-co-linear, or conversely, would have been exposed to one condition and then to the other in some unknown proportion and could not reliably be used to test the hypothesis. Using only those subjects who were co-linear with the teacher at both pretest and posttest, or non-co-linear at both, reduced the number of subjects by 34 of the 76, leaving 42. At this level of "purity," so to speak, I judged that the co-linearity index should be effective enough to separate sheep from goats.

(5) Hypothesis 2 specifies a further step in selection. If the co-linearity index is applied to the subject's own two responses, one at pretest and one at posttest, non-co-linearity "pre-to-post" would imply that the subject has changed his viewpoint in the interim, even though the viewpoints at both times are co-linear (or non-co-linear) with that of the teacher. Of the 42 students used in testing Hypothesis 1, 6 gave non-co-linear pre-to-post responses, leaving 36 subjects in the test of Hypothesis 2.

Quiz Grade z-Scores

Co-linearity, then, applied in the manner described, is the independent variable for Hypotheses 1 and 2. The dependent variable is the mean grade made by the subject on quizzes written and graded by his teacher. Within each of the seven classes, each quiz was given equal weight in the total score. In order to compare quiz grades across classes, z-scores were then

TABLE 2

Difference between Z-scores on Quizzes for Students Co-linear and Non-co-linear with the Instructor

	N	Z-scores		
		Mean	Range	S.D.
Co-linear with instructor	21	.51	-1.16 to 2.77	1.18
Non-co-linear with instructor	21	-.15	-2.56 to 1.74	1.18

TABLE 3

Difference between Z-scores on Quizzes for Students Co-linear and Non-co-linear with the Instructor, for Students Who Were Co-linear with Themselves from Pretest to Posttest

	N	Z-scores		
		Mean	Range	S.D.
Co-linear with instructor	17	.60	-1.16 to 2.77	1.38
Non-co-linear with instructor	19	-.25	-2.56 to 1.74	1.17

computed for each class. The z-scores were used as data in all further computations.

RESULTS FOR HYPOTHESIS 1

Dividing the 42 subjects used to test Hypothesis 1 into those co-linear with the instructor at both pretest and posttest (21 subjects) and those non-co-linear with the instructor at both tests (also 21 subjects) the finding is in the proper direction (t-test, satisfies $\alpha > .07$).² Statistics pertinent to this result are shown in Table 2.

Although this result would make acceptance of this hypothesis dubious by itself, it will be seen that this result is entirely consistent with the result of the test of Hypothesis 2, which reaches traditionally acceptable levels of significance.

RESULTS FOR HYPOTHESIS 2

Hypothesis 2 stated that the difference in quiz grades predicted by Hypothesis 1 will be at least as pronounced when only those students are considered whose pretest and posttest rank orders are co-linear.

In the test of Hypothesis 2, there were 17 subjects in the group co-linear with the instructor and 19 in the non-co-linear group. The t-test applied to the quiz scores of these two groups yields a significance level beyond .05. Statistics for this test are shown in Table 3.

² All probability figures in this paper are two-tailed probabilities.

It should be pointed out that the t-test is not entirely appropriate for testing these hypotheses. When the co-linearity index gives a value of non-co-linear, it may be said according to the theory that the subject could not, from any position on the attribute underlying his response, give a rank order of the stimuli which would unfold with that of the other person. But when the index gives a value of co-linear, the converse cannot be said. An index value of "co-linear" indicates only that it *cannot be said*, according to the theory, that the subject's viewpoint is *not* co-linear with that of the other person. It *may or may not* unfold with his. For this reason, a test of co-variation such as product-moment correlation, chi square, or the t-test demands more of the data than can be predicted.

For the reason that a test such as the t-test is treating the data more stringently than the prediction undertakes, the probability of .07 given by the t-test for Hypothesis 1 becomes more acceptable. As was suggested earlier, the result of Hypothesis 1, when compared to that for Hypothesis 2, argues for the correctness of the theoretical derivations, since it was expected on theoretical grounds that the criteria for the co-linearity index used in Hypothesis 2 would give better results than the less stringent criteria used in Hypothesis 1.

TESTS OF ALTERNATIVE HYPOTHESES

It will now be well to raise a few questions about the findings so far given. The first of these has to do with the effect of co-linearity, as contrasted with similar preferences among the stimuli on the part of teacher and student, in predicting mean grades.

Preferences among the Stimuli

It may occur to the reader to wonder whether it might be that the co-linearity index has picked out, among the data, co-linear rank orders which contain the same stimuli in preferred positions. That is, it might be that co-linear persons are those who agree that certain stimuli are *most preferable*. If this were the case, it might be argued that the theoretical derivations were unnecessary, and that agreement with the teacher on quiz answers was foreshadowed by agreement with the teacher on the choice of the most preferred among the five attitude statements.

The tau statistic, which measures rank order similarity, was used as a measure of the degree to which a student and his teacher chose the same stimulus-statements as best. Since the scatter-diagrams suggested that both the tau values and the quiz z-scores were distributed symmetrically, the product-moment correlation was computed between them. The correlation figure was .23 for 34 degrees of freedom, which is far short of a value at which the null hypothesis of no association could reasonably be rejected.

In short, the data fail to give evidence that quiz grades follow a preference for the same stimulus-statements preferred by the teacher. It is not agreement with the teacher on which statements are the best which makes the difference in quiz grades; rather it is judging the statements according to the *same underlying attribute*, regardless of whether the student agrees with the teacher about the most desirable point on the attribute.

Existence of an Attitude Norm

Another possibility which should be examined is that co-linearity with one's particular teacher is not the determining factor, but rather co-linearity with a *normative* ordering of the stimuli. That is, it might not be the interaction of cognitive fields of teacher and student which accounts for the difference in grade-achievement, but rather the sensitivity of the student to a more general "cultural" frame of reference which is merely mediated by the teacher. If this were the case, the data should show a tendency for mutual co-linearity among the responses of the teachers. That this is not the case is shown in Table 4. This table shows that while teachers 1 and 2 are co-linear, and teachers 3 and 4 are co-linear, neither 1 nor 2 is co-linear with either 3 or 4. This indicates that at least two incompatible viewpoints exist among the five teachers. A tendency toward a single viewpoint is not found.

Scholastic Aptitude

In any investigation where symbolic responses are being studied, the possibility can always be entertained that the performance of the subjects may be related to performance on some measure of symbolic skill such as a test of intelligence, scholastic aptitude, or scholastic achievement. If a relation were found between such an ability and the quiz z-score, the novelty of the present findings would be weakened to the extent that co-linearity with the instructor was not independent of the symbolic skill.

TABLE 4

Co-linearity of Viewpoint between Pairs of Teachers

Margins of the table show identification numbers of the teachers. Each cell shows whether the responses of the two indicated teachers are co-linear or non-co-linear.

2	3	4	5	
Co-lin.	Non-co.	Non-co.	Co-lin.	1
	Non-co.	Non-co.	Co-lin.	2
		Co-lin.	Non-co.	3
			Co-lin.	4

The American Council on Education test of scholastic aptitude seemed an appropriate measure with which to examine this possible relationship. A.C.E. scores were available for 100 of the subjects who responded both to pretest and posttest, including 26 of the 36 subjects used in testing Hypothesis 2.

A t-test was carried out to see whether the co-linearity index somehow selected groups which differed in A.C.E. scores. No difference was demonstrable between the group co-linear with the teacher and the group non-co-linear with the teacher in regard to mean A.C.E. score.

One would conclude from this result that members of the co-linear group were drawn from the same level of A.C.E. scores as members of the non-co-linear group. The difference in z-scores between the two groups can be attributed to the co-linearity condition and not to any difference in scholastic aptitude. To check whether scholastic aptitude could in any case have differentiated among quiz grades, A.C.E. scores for the 100 available cases were correlated with quiz grades, and a positive correlation of .42 was found, which is significant beyond the .05 level. The nonsignificant result of the t-test of the 26 cases in the co-linearity groups, nevertheless, argues that the co-linearity effect on quiz grades was not due to differential selection of scholastic ability.

In sum, it seems clear that the co-linearity index predicts a difference among quiz grades which is not attributable to the relation between A.C.E. scores and quiz grades, to response to an attribute norm, or to a preference for the stimulus-statements preferred by the teacher.

DISCUSSION

This research examines the relation between an interaction variable (co-linearity) and an individual variable (success in choosing "correct" answers in quizzes). It examines the effect on individual responses when a certain kind of "meshing" exists or does not exist between the viewpoints which shape the communicative acts of two or more persons.

In describing the intent and implications of this research it is important to make clear the structural difference between an interaction variable such as co-linearity (or deviation from a group norm, or certain kinds of power relations, to give other examples) and experimental variables which are constructed by comparing individuals on some specific attribute. Examples of the latter would be found in selecting pairs or groups of persons according to how they compared with each other on liking for dancing, years of education, authoritarianism, or some other particular attribute. A variable constructed in this fashion enables hypotheses to be made of this type: if individuals who are high (or low, or the same, or different) on attribute x are put in communication with each other, then these individuals will show certain behaviors rather than others.

An hypothesis concerning an interaction variable such as co-linearity has a different form: if individuals communicate on the basis of the same attribute x , regardless of what attribute it may be and regardless of whether the individuals are high, low, the same, or different on this attribute, then they will show certain behaviors rather than others.

In ascertaining co-linearity between pairs of persons in the present research no attempt was made to put a label on the attribute which two co-linear individuals used in common in organizing their responses to the stimuli. Nor was any attempt made to find out whether one co-linear teacher-pupil pair was responding according to the same attribute utilized by another co-linear teacher-pupil pair. The "content" of the attribute mediating the responses of any pair was unknown. Further, there was good evidence that some co-linear pairs must have been judging the stimuli on the basis of different attributes from other pairs, as indicated by Table 4. The predictions and analyses of this research were made with regard only to the *structural* similarity of orientations toward the stimuli as reflected in the co-linearity index, and without regard to the content character of the orientations, indeed without any knowledge of their content character.

Many investigations of social process require the study of communication, of the transmission of information, of the "understanding" of stimuli presented by some persons to others. The effects of this communication depend on the abilities of the individuals involved, and often upon the agreement or disagreement which exists between the communicators in regard to the content of the communication. But beyond these factors the present research argues that whatever the abilities of the individuals and whatever the extent of their agreement, the effects of the communication also depend on the structural similarity of the viewpoints which mediate "sending" and "receiving" in the communicative interaction.

SUMMARY

This paper has been concerned with the proposition that similarity of structure between two cognitive fields increases the efficacy of communication between them.

Communication or interaction between two persons is conceived as being mediated by the cognitive structure, or space of potential responses, associated with each individual. The response spaces, in turn, can be described in terms of the attributes which mediate the responses; that is, the attributes of the objects in respect to which the individual makes his responses. Further, an individual may combine the attributes mediating his responses into one composite attribute. This composite attribute is compounded in different ways by different individuals, and such a composite attribute underlying the responses of an individual permits some responses and precludes others. I have described criteria for deciding whether or not a com-

posite attribute can be inferred reliably to underlie the responses, and if so, whether or not the composite attributes being used by two communicating individuals might be permitting or precluding the same responses.

If responses of two individuals can be interpreted to be mediated by the same underlying attribute, the responses are termed "co-linear," and "non-co-linear" if they cannot. Computation of an index of co-linearity rests upon the "unfolding technique" of Coombs.

In these terms, we can index similarity of cognitive structure with the co-linearity index and put into more precise terms the proposition with which we began; namely, where changes in orientations occur as a result of communication, the changes will be more pronounced for co-linear communicating pairs and less pronounced for non-co-linear pairs.

This hypothesis was tested by presenting statements to teachers and students concerning the introductory course in psychology in which they interacted for the period of a semester. It was predicted that students co-linear with the teacher would get higher grades on quizzes than students non-co-linear with the teacher. This prediction was well supported by the results. Further examinations of the data provided evidence that the higher grades on the part of co-linear students could not be accounted for by differences in scholastic ability as measured by A.C.E. scores, nor by conformity to a common attitude norm, nor by a preference for the same attitude position as that held by the teacher.

The results of this research imply that differences in abilities between communicators and differences in agreement concerning the content of communication must fail to account for certain effects of communication which can be accounted for by similarity of cognitive structure between the communicators.

Manuscript received: March 5, 1956

Philip J. Runkel
Bureau of Educational Research
1007 S. Wright Street
Champaign, Illinois

REFERENCES

1. Coombs, C. H., "Psychological Scaling without a Unit of Measurement," *Psychological Review*, 1950, 57, 145-158.
2. Coombs, C. H., "Theory and Methods of Social Measurement," in L. Festinger and D. Katz (eds.), *Research Methods in the Behavioral Sciences*. New York: Dryden, 1953.
3. Coombs, C. H., and R. C. Kao, *Non-metric Factor Analysis*, Ann Arbor, Michigan: University of Michigan Press, Engineering Research Bulletin No. 38, 1955.
4. Newcomb, T. M., "An Approach to the Study of Communicative Acts," *Psychological Review*, 1953, 60, 393-404.
5. Runkel, P. J., "Cognitive Facilitation of Communication Effects," Unpublished doctoral dissertation, University of Michigan, 1956.

The Role of Test Intelligence and Occupational Background as Factors in Occupational Choice

C. T. PIHLBLAD, *University of Missouri*

C. L. GREGORY, *University of Missouri*

The purpose of this paper is an attempt to evaluate the relative influence of ability, as measured by score on a standard intelligence test, and occupational background, as shown by parents' occupation, in influencing the selection of an occupation. The material for the investigation has already been described more completely in earlier studies (8). Briefly, the data consist of the following: (a) test scores on the Ohio Psychological Test for 1770 male students in the senior class of small-town Missouri high schools in 116 communities during the years 1939-1940, (b) information concerning parental occupation for each subject at the time of school attendance, (c) occupations followed by the subjects in 1950-1952, ten to twelve years after completing high school. The test score data were supplied by W. R. Carter, Director of the Missouri University College Aptitude Testing Program. Data concerning the current occupation of the subject were supplied by local informants, largely former teachers, school officials, parents, and relatives.

The relationship between test intelligence and occupation has been the subject of a number of earlier investigations (1, 2, 4, 5). Most of these studies have been limited to a comparison of intelligence of school children classified on the basis of the parental occupation. A few studies have correlated test scores from school records with occupations later followed by subjects, although the difficulty of obtaining such data have made them rather rare (3). The results of both types of studies indicate a rather close relationship between occupation and test score, with highest scores appearing among those in the professions and clerical occupations and the lowest scores among unskilled workers and farmers.

In our paper on intelligence as a factor in occupational choice (7) we showed that there was distinct tendency for test intelligence to be related both to the subject's own choice of occupation as well as to the occupation of his parent. Test scores averaged significantly higher for those who entered the professions, clerical occupations, and business than they did for those who became farmers and manual workers. Likewise, the scores for sons of professional and business-class fathers averaged higher than did the scores for sons of farmers and manual workers. Finally, we have shown that there was a distinct and disproportionate tendency for subjects to enter the same or similar occupations to those in which their fathers were

employed at the time of the subject's school attendance (6). Thus all three factors, test intelligence, occupational background, and choice of occupation, are interrelated. It will be the purpose of this paper to analyze this interrelationship further.

The basic question is the degree to which test score is separate and independent from occupational background as a factor in influencing occupational choice. That is, do high-score individuals in low-status occupations tend to gravitate toward higher status occupations, and, conversely, do low-score individuals in higher status occupational levels tend to drop down in the occupational scale?

That the intelligence factor may be directly and independently related to occupational selection is suggested in a comparison of the mean test score for each occupational level when the subjects are classified on (a) their own occupation followed a decade after leaving high school and (b) their occupational background as judged by their father's occupation at the time of the student's school attendance. These data are shown in Table 1.

Comparison of the two series of means shows a definite relation between test score and occupation, whether the classification is based on the sub-

TABLE 1
*Mean Ohio Test Scores for Male Subjects Classified by Own Occupation and Father's Occupation**

Occupational Level	Subject's Own Occupation		Father's Occupation	
	Number	Mean Score	Number	Mean Score
Total	1770†	48.1	1741†	48.8
Professional other than teaching	162	66.0	62	64.4
Teachers	80	61.3	15	64.3
Students	43	57.6	—	—
Military service	135	55.4	—	—
Clerical	58	50.9	21	55.0
Business and sales	347	47.6	209	51.2
Skilled workers	232	46.1	147	50.9
Semiskilled and unskilled workers	311	43.4	230	50.0
Farmers	402	43.2	1057	46.6

* The standard error of the differences between all possible pairings of these means ranged from 1 to 5 units on the Ohio Scale. Most of the errors were grouped around 2.5 to 3.0 units, so that any absolute difference of 5 or 6 points or larger on the Ohio Scale is generally statistically significant at the 5 per cent level.

† The number of cases in the two classifications differs since information concerning father's occupation was not available for all cases for whom the subject's own occupation was reported.

ject's own occupation or on that of the parent. Persons who have entered the professions, those who were classified as students, or those in the military services showed significantly higher scores than did those who were engaged in farming or other manual occupations. Similarly, subjects from white-collar homes, on the average, made higher scores than did those with home backgrounds in manual labor or in farming (8). It may be noted, however, that the range of scores when the classification is based on the subject's own occupation is somewhat greater than when based on the father's occupation; in the former from a high of 66 for professionals to a low of 43 among farmers, a range of 23 points on the Ohio Scale. Based on the father's occupation the range is from 64 for teachers and other professionals to 47 for farmers, or 17 points. This would seem to suggest that the brighter sons of farmers and workers tend to gravitate upward in the occupational scale and that duller sons from white-collar homes tend to move downward. Such an hypothesis seems to be strengthened when we compare separately the means for each occupational level between the two classifications. For example, classified by the father's occupation, farmers' sons have a mean score of 47, but our farmer subjects averaged only 43. The inference can be drawn that sons with higher test scores tended to move out of farming. Sons of unskilled workers averaged 50, but subjects occupied in unskilled work scored a mean of 43. Again there

TABLE 2
*Mean Ohio Test Scores for 1479 Sons Classified by Father's Occupation and Subject's Own Classification**

Father's Occupation	(1) Same Occupation as Father		(2) Shifting to High-status Occupation†		(3) Shifting to Low-status Occupation‡		(4) Total	
	No.	Mean	No.	Mean	No.	Mean	No.	Mean
Professional	24	68.4	23	68.0	4	52.3	51	67.0
Teachers	3	65.3	9	68.6	2	56.0	14	66.1
Clerical workers	0	00.0	10	63.7	10	43.5	20	53.6
Salespeople and proprietors	88	51.3	60	58.3	41	43.5	189	51.8
Skilled workers	42	49.7	56	52.2	20	41.0	118	51.4
Semi- and unskilled workers	59	44.2	98	51.4	33	48.7	190	48.7
Farmers	361	42.9	286	52.4	250	43.1	897	46.0
Total	577	46.0	542	54.0	360	43.7	1479	48.4

* Excludes 291 cases where occupational classification of both son and father was impossible.

† High Status: Professional, Teachers, Clerical Workers, Military Service, Student, and Salespeople and Proprietors.

‡ Low Status: Semi-skilled and Unskilled, Skilled, and Farmers.

seems to be some selection. At the other end of the scale downward selection seems to be indicated by the fact that persons engaged in the professions had a somewhat higher score than did the sons of professional fathers.

The hypothesis that intelligence plays an independent role in the process of occupational shifting and inheritance can be tested in another way, that is, by comparing the mean scores for those who have remained in the paternal occupation and those who have shifted away from it. Table 2 presents such a classification. In the table we have grouped our subjects into two divisions, "high-" and "low-" status occupations. In the first category we have placed largely the white-collar occupations—teachers, other professionals, clerical workers, sales people and proprietors, military service, and students. In the low-status group we have classified those engaged in manual labor, personal service, and farming. The classification is not entirely satisfactory, nor are the distinctions very sharp. Probably many skilled workers are closer in both income and social status to many salespeople and clerical workers than they are to unskilled workers. The farmer category includes large-scale farmers, subsistence farmers in the Ozarks, and sharecroppers in southeast Missouri. Nevertheless, in general, workers in the high-status class are occupied in "clean" work, requiring more formal training and experience, and are accorded higher remuneration and social status than are those in the low-status group. To make possible a comparison of the mean scores for those who have remained in the father's occupation with those who have shifted out of it, we have classified the subjects into three groups: (a) those who have remained at the same occupational class as the father (col. 1), (b) those who have shifted into a high-status occupation (col. 2), and (c) those who have shifted into a low-status occupation (col. 3). In column 4 are shown the mean scores for the total number in each occupational class.

If we compare first the mean score for the total of all occupational categories combined with each of the three classes—those who remained in the father's occupation, those who shifted to a high-status occupation, and those who shifted to a low-status employment—it is apparent that significant differences appear. The highest mean [54] appears among those who shifted to a high-status occupation and the lowest among those who shifted to low-status work [44].

The selective process appears even clearer if we compare separately the mean score in each of the three groups (columns 1, 2, and 3) for each of the occupational levels. While many of these comparisons are of doubtful statistical validity, in light of the limited number of cases in each occupational class, there is a certain consistency in the results that is suggestive.

T values indicate that for the totals of all occupational classes differences between the means were significant at the 1 per cent level between

(a) those who remained in the father's occupation and those who shifted to a high-status occupation (cols. 1 and 2), (b) those who entered low- and high-status occupations (cols. 2 and 3), (c) the total group (col. 4), and all three other classes (1, 2, and 3). Between the means of those who remained in the father's occupation and those who entered low-status employment differences were not significant.

If we deal separately with each occupational level, the limited number of cases makes it more difficult to establish significance between the mean scores. Nevertheless, if we compare high- and low-status categories, *T* values indicate significance at the 1 per cent level for farmers and sales-people, at the 5 per cent level for skilled workers. Other differences are not significant. Between those in the same occupation as the father (col. 1) and those shifting to a high-status occupation (col. 2) differences are significant at the 5 per cent level only for unskilled and farmers. Between the total group (col. 4) and those in the father's occupation (col. 1) differences are significant at the 1 per cent level only for the farmers. Other differences are not significant.

In the professional group those who remained in the same occupational level as the father had the same mean score as those who shifted into another high-status occupation [68]. But those who shifted out of this level toward a low-status occupation showed a mean score 16 points lower [52]. None of the 20 sons of clerical workers followed clerical occupations. Ten entered a high-status occupation with a mean score of 64, while ten dropped into low-status employment with an average of 44. Sons from business-class homes (sales and proprietary) who remained at this occupational level averaged 51. Those who shifted into some other high-status work averaged 58 while those who dropped to a low-status occupation had a mean of 44. Among skilled workers those who shifted to a low-status occupation had a markedly lower score than did either those who remained in skilled work or those who shifted to high-status work. Among farmers the selective process is clearer and statistically more significant than for other groups. The lowest mean score [44] appears for those who left farming and entered other low-status employment, that is, manual labor. Those who shifted out of farming into white-collar employment scored highest [54], while an intermediate mean of 46 was made by those who remained in farming.

While the differences between the means of those who have remained in the parental occupation and those who have shifted out of it, either up or down, are not always large and the number of cases often too small to establish statistical validity, their consistency seems to support our hypothesis, namely, that occupational differentiation is selective of the kinds of

abilities which are measured by an intelligence test, independent of the occupational backgrounds of the subject. We may put it more plainly. The higher scores of those in the upper status occupations are partly due to the fact that they recruit from occupational backgrounds which disproportionately produce children with high average scores but partly, also, because they recruit disproportionately from the superior group in lower status occupations. Conversely, also, there seems to be a tendency for low-score individuals from upper occupational backgrounds to find their own occupational niche at a somewhat lower level than that in which their fathers were occupied.

THE RELATIVE INFLUENCE ON OCCUPATIONAL CHOICE OF OCCUPATIONAL BACKGROUND AND INTELLIGENCE

In this final section of the paper we would like to push our analysis one step further, namely, to attempt to measure the relative influence of each of our two factors, occupational background and test intelligence, on the choice of the subject's own occupation. In Table 3 are presented data which throw some light on this question. Here we have divided our subjects, on the basis of both their own and their father's occupation, into the two categories, high and low status, which were used earlier. In addition we have also classified the subjects into two test-score categories, a high group with scores of 50 and above and a low group with scores below 50. The

TABLE 3

Differences in Proportions of Subjects Selecting High-status Occupations with Variations in Ohio Score and Father's Occupation

Subject's Test Score	Father's Occupation	Subject's Occupation High Status	Difference
1. 50 and above	High status	88.9	20.0
Below 50	High status	68.9	
2. 50 and above	Low status	48.3	17.4
Below 50	Low status	30.9	
3. 50 and above	High status	88.9	40.6
50 and above	Low status	48.3	
4. Below 50	High status	68.9	38.0
Below 50	Low status	30.9	

High-status Occupation: Professional and teachers, clerical worker, military science, student, and sales-people and proprietors.

Low-status Occupation: Farmers, unskilled, skilled, and personal service.

per cent of the total group in each of these classifications is shown. By varying one factor at a time and noting the differences produced by such variation, we may get a rough measure of the relative effect of the two variables on the choice of occupation.

In lines 1 and 2 we have held constant the factor of father's occupation while varying the test score. We may note that 89 per cent of the subjects with high-status occupational background who scored 50 and above entered a high-status occupation while only 69 per cent who scored below 50 entered such an occupation. Of the sons with low-status occupational backgrounds 48 per cent entered a high-status occupation when they scored 50 and over while only 31 per cent when they scored below 50. Thus we get a variation of approximately twice the variation produced by shifting the test score. It seems reasonable, then, to conclude that a father's occupation exercises about twice the influence of test intelligence on occupational choice.

The results of this analysis would seem to square with the common observation that a great many factors, other than ability, influence occupational choice and the attainment of positions on the occupational ladder which carry both prestige and material reward. The results are consistent with the common observation that mediocre sons of fathers in high-status occupations will often attain relatively favored positions themselves, probably as a consequence of the advantages of family background. The mediocre son of a physician may not be able to qualify for medicine, but he is unlikely to become a manual worker. He may have to settle for a position in business, teaching, or clerical work. He will probably not have to, or want to take a job even as an electrician or welder to say nothing of becoming a ditchdigger. Similarly, the bright son of a manual worker or farmer will find it difficult to overcome economic and social obstacles to the attainment of a niche in medicine, law, and science, or the upper echelons of business and industry; but he will attain such objectives more frequently than the duller subject from a low-status background.

In final conclusion a note of caution should be injected in the interpretation of our results. It is doubtful that we can draw any sweeping inference to the effect that our society is moving toward greater rigidity or toward a closed-status system. While parental occupation seems to have more influence on occupational choice than does intelligence, yet, as we demonstrated, intelligence is also independent of occupation in its influence on choice of calling. Brighter individuals from all walks do find their way into the upper status occupations with more than chance frequency. We must also keep in mind that our data pertain only to small-town and rural residents in Middle Western communities. The selective process may well be different in urban situations. Finally, it should be clear that intelligence tests measure a rather narrow set of abilities which

probably reveal something about the subjects' capacity for achievement in school tasks but are questionable measures of other aspects of social worth.

Manuscript received: January 31, 1956

Revised manuscript received: April 13, 1956

C. T. Pihlblad

Department of Sociology and Anthropology

University of Missouri

Columbia, Missouri

REFERENCES

1. Anderson, A., J. C. Brown, and M. J. Bowman, "Intelligence and Occupational Mobility," *Journal of Political Economy*, 1952, 60, 218-239.
2. Foulds, G. A., and J. C. Ranen, "Intellectual Ability and Occupational Grade," *Occupational Psychology*, 1948, 22, 197-203.
3. Gist, N. P., C. T. Pihlblad, and C. L. Gregory, *Selective Factors in Occupation and Migration*, Columbia, Mo.: University of Missouri, 1943.
4. Lorimer, F., and F. Osborn, *Dynamics of Population: Social and Biological Significance of Changing Birth Rates in the United States*, New York: Macmillan, 1934.
5. Moser, W. E., "Vocational Preference as Related to Mental Ability," *Occupations*, 1949, 27, 460-461.
6. Pihlblad, C. T., and C. L. Gregory, "Occupational Mobility in Small Communities in Missouri," *Rural Sociology*, in press.
7. Pihlblad, C. T., and C. L. Gregory, "Occupational Selection and Intelligence in Rural Communities and Small Towns in Missouri," *American Sociological Review*, 1956, 21, 63-71.
8. Pihlblad, C. T., and C. L. Gregory, "Selective Aspects of Migration Among Missouri High School Graduates," *American Sociological Review*, 1954, 15, 314-324.